PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12N 9/28 — A1

(11) International Publication Number: WO 00/29560

(43) International Publication Date: 25 May 2000 (25.05.00)

(21) International Application Number: PCT/DK99/00628

(22) International Filing Date: 16 November 1999 (16.11.99)

(30) Priority Data:
PA 1998 01495 16 November 1998 (16.11.98) DK

(71) Applicant: NOVO NORDISK A/S [DK/DK]; Novo Allé, DK-2880 Bagsværd (DK).

(72) Inventors: SVENDSEN, Allan; Novo Nordisk A/S, Novo Allé, DK-2880 Bagsværd (DK). KJÆRULFF, Søren; Novo Nordisk A/S, Novo Allé, DK-2880 Bagsværd (DK). BISGÅRD-FRANTZEN, Henrik; Novo Nordisk A/S, Novo Allé, DK-2880 Bagsværd (DK). ANDERSEN, Carsten; Novo Nordisk A/S, Novo Allé, DK-2880 Bagsværd (DK).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: α-AMYLASE VARIANTS

(57) Abstract

The invention relates to a variant of a parent Termamyl-like α-amylase, comprising mutations in two, three, four, five or six regions/positions. The variants have increased stability at high temperatures (relative to the parent). The invention also relates to a DNA construct comprising a DNA sequence encoding an α-amylase variant of the invention, a recombinant expression vector which carries a DNA construct of the invention, a cell which is transformed with a DNA construct of the invention, the use of an α-amylase variant of the invention for washing and/or dishwashing, textile desizing, starch liquefaction, a detergent additive comprising an α-amylase variant of the invention, a manual or automatic dishwashing detergent composition comprising an α-amylase variant of the invention, a method for generating a variant of a parent Termamyl-like α-amylase, which variant exhibits increased.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

Title: α-amylase variants

FIELD OF THE INVENTION

The present invention relates to novel variants of parent Termamyl-like α-amylases with altered properties relative of the parent alpha-amylase. Said properties include increased stability, e.g., at acidic pH, e.g., at low calcium concentrations and/or high temperatures. Suach variants are suitable for a number of applications, in particular, industrial starch processing (*e.g.*, starch liquefaction or saccharification).

BACKGROUND OF THE INVENTION

α-Amylases (α-1,4-glucan-4-glucanohydrolases, EC 3.2.1.1) constitute a group of enzymes which catalyze hydrolysis of starch and other linear and branched 1,4-glucosidic oligo- and polysaccharides.

There is a very extensive body of patent and scientific literature relating to this industrially very important class of enzymes. A number of α-amylase such as Termamyl-like α-amylases variants are known from, *e.g.*, WO 90/11352, WO 95/10603, WO 95/26397, WO 96/23873 and WO 96/23874.

WO 96/23874 provides the three-dimensional, X-ray crystal structural data for a Termamyl-like α-amylase which consists of the 300 N-terminal amino acid residues of the *B. amyloliquefaciens* α-amylase and amino acids 301-483 of the C-terminal end of the *B. licheniformis* α-amylase comprising the amino acid sequence (the latter being available commercially under the tradename Termamyl™), and which is thus closely related to the industrially important *Bacillus* α-amylases (which in the present context are embraced within the meaning of the term "Termamyl-like α-amylases", and which include, *inter alia*, the *B. licheniformis*, *B. amyloliquefaciens* and *B. stearothermophilus* α-amylases). WO 96/23874 further describes methodology for designing, on the basis of an analysis of the

structure of a parent Termamyl-like α-amylase, variants of the parent Termamyl-like α-amylase which exhibit altered properties relative to the parent.

BRIEF DISCLOSURE OF THE INVENTION

The present invention relates to novel α-amylolytic variants (mutants) of a Termamyl-like α-amylase, in particular variants exhibiting increased stability at acidic pH at high temperatures (relative to the parent) which are advantageous in connection with, e.g., the industrial processing of starch (starch liquefaction, saccharification and the like) as described in US Patent No. 3,912,590 and EP patent publications Nos. 252,730 and 63,909.

Starch conversion

A "traditional" starch conversion process degrading starch to lower molecular weight carbohydrate components such as sugars or fat replacers includes a debranching step.

"Starch to sugar" conversion

In the case of converting starch into a sugar the starch is depolymerized. A such depolymerization process consists of a pretreatment step and two or three consecutive process steps, viz. a liquefaction process, a saccharification process and dependent on the desired end product optionally an isomerization process.

Pre-treatment of native starch

Native starch consists of microscopic granules which are insoluble in water at room temperature. When an aqueous starch slurry is heated, the granules swell and eventually burst, dispersing the starch molecules into the solution. During this "gelatinization" process there is a dramatic increase in viscosity. As the solids level is 30-40% in a typically industrial process, the starch has to be thinned or "liquefied" so that it can be handled. This reduction in viscosity is today

mostly obtained by enzymatic degradation.

Liquefaction

During the liquefaction step, the long chained starch is degraded into branched and linear shorter units (maltodextrins) by an α-amylase (*e.g.*, Termamyl™ SEQ ID NO: 4 herein). The liquefaction process is carried out at 105-110°C for 5 to 10 minutes followed by 1-2 hours at 95°C. The pH lies between 5.5 and 6.2. In order to ensure an optimal enzyme stability under these conditions, 1 mM of calcium is added (40 ppm free calcium ions). After this treatment the liquefied starch will have a "dextrose equivalent" (DE) of 10-15.

Saccharification

After the liquefaction process the maltodextrins are converted into dextrose by addition of a glucoamylase (e.g., AMG™) and a debranching enzyme, such as an isoamylase (US Patent 4,335,208) or a pullulanase (*e.g.*, Promozyme™) (US Patent 4,560,651). Before this step the pH is reduced to a value below 4.5, maintaining the high temperature (above 95°C) to inactivate the liquefying α-amylase to reduce the formation of short oligosaccharide called "panose precursors" which cannot be hydrolyzed properly by the debranching enzyme.

The temperature is lowered to 60°C, and glucoamylase and debranching enzyme are added. The saccharification process proceeds for 24-72 hours.

Normally, when denaturing the α-amylase after the liquefaction step about 0.2-0.5% of the saccharification product is the branched trisaccharide 6^2-α-glucosyl maltose (panose) which cannot be degraded by a pullulanase. If active amylase from the liquefaction step is present during saccharification (i.e., no denaturing), this level can be as high as 1-2%, which is highly undesirable as it lowers the saccharification yield significantly.

Isomerization

When the desired final sugar product is *e.g.* high fructose syrup the dextrose syrup may be converted into fructose. After the saccharification process the pH is increased to a value in the range of 6-8, preferably pH 7.5, and the calcium is removed by ion exchange. The dextrose syrup is then converted into high fructose syrup using, *e.g.*, an immmobilized glucoseisomerase (such as Sweetzyme™).

In the context of the invention the term "acidic pH" means a pH below 7.0, especially below the pH range in which industrial starch liquefaction processes are traditionally performed, as described above, which is between pH 5.5 and 6.2.

In the context of the present invention the term "low Calcium concentration" means concentrations below the normal level used in traditional industrial starch liquefaction processes, such as between 0-40 ppm, preferably between 10-30 ppm, such as between 15-25 ppm Calcium. Normal concentrations vary depending of the concentration of free Ca^{2+} in the corn. Normally a dosage corresponding to 1mM (40ppm) is added which together with the level in corn gives between 40 and 60 ppm free Ca^{2+}.

In the context of the invention the term "high temperature" means temperatures between 95 and 160°C, especially the temperature range in which industrial starch liquefaction processes are normally performed, which is between 95 and 105°C.

The invention further relates to DNA constructs encoding variants of the invention, to methods for preparing variants of the invention, and to the use of variants of the invention, alone or in combination with other α-amylolytic enzymes, in various industrial processes, in particular starch liquefaction.

Nomenclature

In the present description and claims, the conventional one-letter and three-letter codes for amino acid residues are used.

For ease of reference, α-amylase variants of the invention are described by use of the following nomenclature:

Original amino acid(s):position(s):substituted amino acid(s)

According to this nomenclature, for instance the substitution of alanine for asparagine in position 30 is shown as:

Ala30Asn or A30N

a deletion of alanine in the same position is shown as:

Ala30* or A30*

and insertion of an additional amino acid residue, such as lysine, is shown as:

Ala30AlaLys or A30AK

A deletion of a consecutive stretch of amino acid residues, such as amino acid residues 30-33, is indicated as (30-33)* or Δ(A30-N33).

Where a specific α-amylase contains a "deletion" in comparison with other α-amylases and an insertion is made in such a position this is indicated as:

*36Asp or *36D

for insertion of an aspartic acid in position 36

Multiple mutations are separated by plus signs, i.e.:

Ala30Asp + Glu34Ser or A30N+E34S

representing mutations in positions 30 and 34 substituting alanine and glutamic acid for asparagine and serine, respectively. Multiple mutation may also be separated as follows, i.e., meaning the same as the plus sign:

Ala30Asp/Glu34Ser or A30N/E34S

When one or more alternative amino acid residues may be inserted in a given position it is indicated as

A30N,E or

A30N or A30E

Furthermore, when a position suitable for modification is identified herein without any specific modification being suggested, it is to be understood that any amino acid residue may be substituted for the amino acid residue present in the position. Thus, for instance, when a modification of an alanine

in position 30 is mentioned, but not specified, it is to be understood that the alanine may be deleted or substituted for any other amino acid, i.e., any one of: R,N,D,A,C,Q,E,G,H,I,L,K,M,F,P,S,T,W,Y,V.

BRIEF DESCRIPTION OF THE DRAWING

Figure 1 is an alignment of the amino acid sequences of six parent Termamyl-like α-amylases in the context of the invention. The numbers on the Extreme left designate the respective amino acid sequences as follows:

1: SEQ ID NO: 2,
2: amylase
3: SEQ ID NO: 1,
4: SEQ ID NO: 5,
5: SEQ ID NO: 4,
6: SEQ ID NO: 3.

Figure 2 shows the PCR strategy used in Example 1.

DETAILED DISCLOSURE OF THE INVENTION

The Termamyl-like α-amylase

It is well known that a number of α-amylases produced by *Bacillus* spp. are highly homologous on the amino acid level. For instance, the *B. licheniformis* α-amylase comprising the amino acid sequence shown in SEQ ID NO: 4 (commercially available as Termamyl™) has been found to be about 89% homologous with the *B. amyloliquefaciens* α-amylase comprising the amino acid sequence shown in SEQ ID NO: 5 and about 79% homologous with the *B. stearothermophilus* α-amylase comprising the amino acid sequence shown in SEQ ID NO: 3. Further homologous α-amylases include an α-amylase derived from a strain of the *Bacillus* sp. NCIB 12289, NCIB 12512, NCIB 12513 or DSM 9375, all of which are described in detail in WO 95/26397, and the α-amylase described by Tsukamoto et al., Biochemical and Biophysical Research Communications, 151 (1988), pp. 25-31.

Still further homologous α-amylases include the α-amylase produced by the *B. licheniformis* strain described in EP 0252666 (ATCC 27811), and the α-amylases identified in WO 91/00353 and WO 94/18314. Other commercial Termamyl-like *B. licheniformis* α-amylases are Optitherm™ and Takatherm™ (available from Solvay), Maxamyl™ (available from Gist-brocades/Genencor), Spezym AA™ and Spezyme Delta AA™ (available from Genencor), and Keistase™ (available from Daiwa).

Because of the substantial homology found between these α-amylases, they are considered to belong to the same class of α-amylases, namely the class of "Termamyl-like α-amylases".

Accordingly, in the present context, the term "Termamyl-like α-amylase" is intended to indicate an α-amylase which, at the amino acid level, exhibits a substantial homology to Termamyl™, *i.e.*, the *B. licheniformis* α-amylase having the amino acid sequence shown in SEQ ID NO: 4 herein. In other words, a Termamyl-like α-amylase is an α-amylase which has the amino acid sequence shown in SEQ ID NOS: 1, 2, 3, 4, 5, 6, 7 or 8 herein, and the amino acid sequence shown in SEQ ID NO: 1 of WO 95/26397 (the same as the amino acid sequence shown as SEQ ID NO: 7 herein) or in SEQ ID NO: 2 of WO 95/26397 (the same as the amino acid sequence shown as SEQ ID NO: 8 herein) or in Tsukamoto et al., 1988, (which amino acid sequence is shown in SEQ ID NO: 6 herein) or i) which displays at least 60% homology (identity), preferred at least 70%, more preferred at least 75%, even more preferred at least 80%, especially at least 85%, especially preferred at least 90%, especially at least 95%, even especially more preferred at least 97%, especially at least 99% homology with at least one of said amino acid sequences shown in SEQ ID NOS 1: or 2 or 3 or 4 or 5 or 6 or 7 or 8 and/or ii) displays immunological cross-reactivity with an antibody raised against one or more of said α-amylases, and/or iii) is encoded by a DNA sequence which hybridizes, under the low to very high stringency conditions (said conditions described below) to the

DNA sequences encoding the above-specified α-amylases which are apparent from SEQ ID NOS: 9, 10, 11, 12, and 32, respectively, of the present application (which encodes the amino acid sequences shown in SEQ ID NOS: 1, 2, 3, 4, and 5 herein, respectively), from SEQ ID NO: 4 of WO 95/26397 (which DNA sequence, together with the stop codon TAA, is shown in SEQ ID NO: 13 herein and encodes the amino acid sequence shown in SEQ ID NO: 8 herein) and from SEQ ID NO: 5 of WO 95/26397 (shown in SEQ ID NO: 14 herein), respectively.

In connection with property i), the "homology" (identity) may be determined by use of any conventional algorithm, preferably by use of the gap progamme from the GCG package version 8 (August 1994) using default values for gap penalties, i.e., a gap creation penalty of 3.0 and gap extension penalty of 0.1 (Genetic Computer Group (1991) Programme Manual for the GCG Package, version 8, 575 Science Drive, Madison, Wisconsin, USA 53711).

The parent Termamyl-like α-amylase backbone may in an embodiment have an amino acid sequence which has a degree of identity to SEQ ID NO: 4 of at least 65%, preferably at least 70%, preferably at least 75%, more preferably at least 80%, more preferably at least 85%, even more preferably at least about 90%, even more preferably at least 95%, even more preferably at least 97%, and even more preferably at least 99% identity determined as described above

A structural alignment between Termamyl® (SEQ ID NO: 4) and a Termamyl-like α-amylase may be used to identify equivalent/corresponding positions in other Termamyl-like α-amylases. One method of obtaining said structural alignment is to use the Pile Up programme from the GCG package using default values of gap penelties, i.e., a gap creation penalty of 3.0 and gap extension penalty of 0.1. Other structural alignment methods include the hydrophobic cluster analysis (Gaboriaud et al., (1987), FEBS LETTERS 224, pp. 149-155) and reverse threading (Huber, T ; Torda, AE, PROTEIN SCIENCE Vol. 7, No. 1 pp. 142-149 (1998).

For example, the corresponding positions, of target residues found in the C-domain of the *B. licheniformis* α-amylase, in the amino acid sequences of a number of Termamyl-like α-amylases which have already been mentioned are as follows:

Termamyl-like α-amylase					
B. lich. (SEQ ID NO: 4)	S356	Y358	E376	S417	A420
B. amylo. (SEQ ID NO: 5)	S356	Y358	E376	S417	A420
B. stearo.(SEQ ID NO: 3)	----	Y361	----	----	----
*Bac.*WO 95/26397 (SEQ ID NO: 2)	----	Y363	----	S419	----
*Bac.*WO 95/26397 (SEQ ID NO: 1)	----	Y363	----	----	----

As will be described further below mutations of these conserved amino acid residues are very important in relation to increasing the stability at acidic pH and/or at low calcium concentration at high temperatures.

Property ii) (see above) of the α-amylase, *i.e.*, the immunological cross reactivity, may be assayed using an antibody raised against, or reactive with, at least one epitope of the relevant Termamyl-like α-amylase. The antibody, which may either be monoclonal or polyclonal, may be produced by methods known in the art, *e.g.*, as described by Hudson et al., Practical Immunology, Third edition (1989), Blackwell Scientific Publications. The immunological cross-reactivity may be determined using assays known in the art, examples of which are Western Blotting or radial immunodiffusion assay, *e.g.*, as described by Hudson et al., 1989. In this respect, immunological cross-reactivity between the α-amylases having the amino acid sequences SEQ ID NOS: 1, 2, 3, 4, 5, 6, 7, or 8 respectively, have been found.

The oligonucleotide probe used in the characterization of the Termamyl-like α-amylase in accordance with property iii)

above may suitably be prepared on the basis of the full or partial nucleotide or amino acid sequence of the α-amylase in question.

Suitable conditions for testing hybridization involve presoaking in 5xSSC and prehybridizing for 1 hour at ~40°C in a solution of 20% formamide, 5xDenhardt's solution, 50mM sodium phosphate, pH 6.8, and 50mg of denatured sonicated calf thymus DNA, followed by hybridization in the same solution supplemented with 100mM ATP for 18 hours at ~40°C, followed by three times washing of the filter in 2xSSC, 0.2% SDS at 40°C for 30 minutes (low stringency), preferred at 50°C (medium stringency), more preferably at 65°C (high stringency), even more preferably at ~75°C (very high stringency). More details about the hybridization method can be found in Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor, 1989.

In the present context, "derived from" is intended not only to indicate an α-amylase produced or producible by a strain of the organism in question, but also an α-amylase encoded by a DNA sequence isolated from such strain and produced in a host organism transformed with said DNA sequence. Finally, the term is intended to indicate an α-amylase which is encoded by a DNA sequence of synthetic and/or cDNA origin and which has the identifying characteristics of the α-amylase in question. The term is also intended to indicate that the parent α-amylase may be a variant of a naturally occurring α-amylase, i.e., a variant which is the result of a modification (insertion, substitution, deletion) of one or more amino acid residues of the naturally occurring α-amylase.

Parent hybrid α-amylases

The parent α-amylase (backbone) may be a hybrid α-amylase, i.e., an α-amylase which comprises a combination of partial amino acid sequences derived from at least two α-amylases.

The parent hybrid α-amylase may be one which on the basis of

amino acid homology and/or immunological cross-reactivity and/or DNA hybridization (as defined above) can be determined to belong to the Termamyl-like α-amylase family. In this case, the hybrid α-amylase is typically composed of at least one part of a Termamyl-like α-amylase and part(s) of one or more other α-amylases selected from Termamyl-like α-amylases or non-Termamyl-like α-amylases of microbial (bacterial or fungal) and/or mammalian origin.

Thus, the parent hybrid α-amylase may comprise a combination of partial amino acid sequences deriving from at least two Termamyl-like α-amylases, or from at least one Termamyl-like and at least one non-Termamyl-like bacterial α-amylase, or from at least one Termamyl-like and at least one fungal α-amylase. The Termamyl-like α-amylase from which a partial amino acid sequence derives may, *e.g.*, be any of those specific Termamyl-like αa-mylase referred to herein.

For instance, the parent α-amylase may comprise a C-terminal part of an α-amylase derived from a strain of *B. licheniformis*, and a N-terminal part of an α-amylase derived from a strain of *B. amyloliquefaciens* or from a strain of *B. stearothermophilus*. For instance, the parent α-amylase may comprise at least 430 amino acid residues of the C-terminal part of the *B. licheniformis* α-amylase. A such hybrid Termamyl-like α-amylase may be identical to the *Bacillus licheniformis* α-amylase shown in SEQ ID NO: 4, except that the N-terminal 35 amino acid residues (of the mature protein) is replaced with the N-terminal 33 amino acid residues of the mature protein of the *Bacillus amyloliquefaciens* α-amylase (BAN) shown in SEQ ID NO: 5. A such hybrid may also consist of an amino acid segment corresponding to the 68 N-terminal amino acid residues of the *B. stearothermophilus* α-amylase having the amino acid sequence shown in SEQ ID NO: 3 and an amino acid segment corresponding to the 415 C-terminal amino acid residues of the *B. licheniformis* α-amylase

having the amino acid sequence shown in SEQ ID NO: 4.

The non-Termamyl-like α-amylase may, *e.g.*, be a fungal α-amylase, a mammalian or a plant α-amylase or a bacterial α-amylase (different from a Termamyl-like α-amylase). Specific examples of such α-amylases include the *Aspergillus oryzae* TAKA α-amylase, the *A. niger* acid α-amylase, the *Bacillus subtilis* α-amylase, the porcine pancreatic α-amylase and a barley α-amylase. All of these α-amylases have elucidated structures which are markedly different from the structure of a typical Termamyl-like α-amylase as referred to herein.

The fungal α-amylases mentioned above, i.e. derived from *A. niger* and *A. oryzae*, are highly homologous on the amino acid level and generally considered to belong to the same family of α-amylases. The fungal α-amylase derived from *Aspergillus oryzae* is commercially available under the tradename Fungamyl™.

Furthermore, when a particular variant of a Termamyl-like α-amylase (variant of the invention) is referred to - in a conventional manner - by reference to modification (*e.g.*, deletion or substitution) of specific amino acid residues in the amino acid sequence of a specific Termamyl-like α-amylase, it is to be understood that variants of another Termamyl-like α-amylase modified in the equivalent position(s) (as determined from the best possible amino acid sequence alignment between the respective amino acid sequences) are encompassed thereby.

A preferred embodiment of a variant of the invention is one derived from a *B. licheniformis* α-amylase (as parent Termamyl-like α-amylase), e.g., one of those referred to above, such as the *B. licheniformis* α-amylase having the amino acid sequence shown in SEQ ID NO: 4.

<u>Altered properties of variants of the invention</u>

The following discusses the relationship between alterations/mutations which may be present in variants of the invention, and desirable alterations in properties (relative to

those a parent, Termamyl-like α-amylase) which may result therefrom.

Increased stability at acidic pH and/or low calcium concentration at high temperatures

The present invention relates to a variant of a parent Termamyl-like α-amylase, which variant α-amylase has been altered in comparison to the parent α-amylase in one or more solvent exposed amino acid residues on the surface of the α-amylase to increase the overall hydrophibicity of the α-amylase and/or to increase the overall numbers of methyl groups in the sidechains of said solvent exposed amino acid residues on the surface.

In a preferred embodiment one or more solvent exposed amino acid residues on a concav surface with inwards bend are altered to more hydrophobic amino acid residues.

In another preferred embodiment one or more solvent exposed amino acid residues on a convex surface are altered to increase the number of methyl groups in the sidechain.

The present invention relates to an α-amylase variant of a parent Termamyl-like α-amylase, comprising an alteration at one or more positions selected from the group of:

E376, S417, A420, S356, Y358;

wherein (a) the alteration(s) are independently

(i) an insertion of an amino acid downstream of the amino acid which occupies the position,

(ii) a deletion of the amino acid which occupies the position, or

(iii) a substitution of the amino acid which occupies the position with a different amino acid,

(b) the variant has α-amylase activity and (c) each position corresponds to a position of the amino acid sequence of the parent Termamyl-like α-amylase having the amino acid sequence of SEQ ID NO: 4.

In an embodiment the alteration is one of the following

substitutions:
E376A,R,D,C,Q,G,H,I,K,L,M,N,F,P,S,T,W,Y,V.
In a preferred embodiment the substitution is: E376K.

In an embodiment the alteration is one of the following substitutions: S417A,R,D,C,E,Q,G,H,I,K,L,M,N,F,P,T,W,Y,V;
In a preferred embodiment the substitution is S417T.

In an embodiment the alteration is one of the following substitutions A420R,D,C,E,Q,G,H,I,K,L,M,N,F,P,S,T,W,Y,V;
In a preferred embodiment the substitution is: A420Q,R.

In an embodiment the alteration is one of the following substitutions: S356A,R,D,C,E,Q,G,H,I,K,L,M,N,F,P,T,W,Y,V.

In an embodiment the alteration is one of the following substitutions Y358A,R,D,C,E,Q,G,H,I,K,L,M,N,F,P,S,T,W,V.
In a preferred embodiment the substitution is Y358F.

In an embodiment of the invention a variant comprises one or more of the following substitutions: E376K, S417T, A420Q,R, S356A, Y358F.

The increase in stability at acidic pH and/or low calcium concentration at high temperatures may be determined using the method described below in Example 2 illustrating the invention.

The parent Termamyl-like α-amylase used as the backbone for preparing variants of the invention may be any Termamyl-like α-amylases as defined above.

Specifically contemplated are parent Termamyl-like α-amylases selected from the group derived from *B. licheniformis*, such as *B. licheniformis* strain ATCC 27811, *B. amyloliquefaciens*, *B. stearothermophilus*, *Bacillus* sp. NCIB 12289, NCIB 12512, NCIB 12513 or DSM 9375, and the parent Termamyl-like α-amylases depicted in SEQ ID NOS: 1, 2, 3, 4, 5, 6, 7 and 8.

In an embodiment of the invention the parent Termamyl-like α-amylase is a hybrid α-amylase being identical to the *Bacillus licheniformis* α-amylase shown in SEQ ID NO: 4 (Termamyl), except that the N-terminal 35 amino acid residues (of the mature protein) is replaced with the N-terminal 33 amino acid residues of the mature protein of the *Bacillus amyloliquefaciens* α-amylase (BAN) shown in SEQ ID NO: 5. The parent Termamyl-like

hybrid α-amylase may be the above mentioned hybrid Termamyl-like α-amylase which further has the following mutations: H156Y+181T+190F+209V+264S (using the numbering in SEQ ID NO: 4). Said backbone is referred to below as "LE174".

The parent α-amylase may advantageously further have a mutation in one or more of the following positions: K176, I201 and H205 (using the numbering in SEQ ID NO: 4), especially one or more the following substitutions: K176R, I201F, and H205N (using the numbering in SEQ ID NO: 4), such as specifically the following substitutions: K176R+I201F+H205N (using the numbering in SEQ ID NO: 4).

The inventors have found that the above mentioned variants have increased stability at pHs below 7.0 (i.e., acidic pH) and/or at calcium concentration below 1mM (40ppm) (i.e, low calcium concentrations) at temperatures in the range from 95 to 160°C (i.e., high temperatures) relative to the parent Termamyl-like α-amylase.

Alterations (e.g., by substitution) of one or more solvent exposed amino acid residues which 1) increase the overall hydrophobicity of the enzyme, or 2) increase the number of methyl groups in the sidechains of the solvent exposed amino acid residues improve the temperature stability. It is preferred to alter (e.g., by substitution) to more hydrophobic residues on a concav surface with inwards bend. On a convex surface alterations (e.g., by substitution) to amino acid residues with an increased number of methyl groups in the sidechain are preferred.

Using the program CAST found on the internet at http://sunrise.cbs.umn.edu/cast/ version 1.0 (release Feb. 1998),(reference: Jie Liang, Herbert Edelsbrunner, and Clare Woodward. 1998. Anatomy of protein Pockets and Cavities: Measurements of binding site geometry and implications for ligand design. Protein Science, 7, pp. 1884-1897), a concave area which access to the surface can be identified. Access to the surface is in the program defined as a probe with a diameter of 1.4Å can pass in and out. Using default parameters in the

CAST program cancave cavities can be found using the Calcium depleted alpha-amylase structure from *B. licheniformis* as found in the Brookhaven database (1BPL):

Three types of interaction can be rationalised:

A. Interaction between the sidechain of the residue and the protein,

B. Interaction between the sidechain of the residue and the surrounding water,

C. Interaction between the water and the protein.

Using the parent Termamyl-like α-amylase shown in SEQ ID NO: 4 as the backbone the following positions are considered to be solvent exposed and may suitably be altered:

E376, S417, A420, S356, Y358.

Corresponding and other solvent exposed positions on the surface of other Termamyl-like α-amylase may be identified using the dssp program by W. Kabsch and C. Sander, Biopolymers 22 (1983) pp. 2577-2637. The convex surfaces can be identified using the the AACAVI program part from the WHATIF package (G. Vriend, Whatif and drug design program. J. Mol. Graph. 8, pp. 52-56. (1990) version 19980317).

In an embodiment of the invention a variant comprises one or more of the following substitutions: E376K, S417T, A420Q,R, S356A, Y358F.

The inventors have found that the stability at acidic pH and/or low calcium concentration at high temperatures may be increased even more by combining mutations in the above mentioned positions, i.e., E376, S417, A420, S356, Y358, (using the SEQ ID NO: 4 numbering) with mutations in one or more of positions K176, I201, and H205.

The following additional substitutions are preferred:

K176A,R,D,C,E,Q,G,H,I,L,M,N,F,P,S,T,W,Y,V;

I201A,R,D,C,E,Q,G,H,L,K,M,N,F,P,S,T,W,Y,V;

H205A,R,D,C,E,Q,G,I,L,K,M,N,F,P,S,T,W,Y,V;

As also shown in Example 2 illustrating the invention combining the following mutations give increased stability:

K176+I201F+H205N+E376K+A420R or

K176+I201F+H205N+S417T+A420Q or K176+I201F+H205N+S356A+Y358F using the hybrid α-amylase referred to as LE174 as the parent Termamyl-like α-amylase.

General mutations in variants of the invention

It may be preferred that a variant of the invention comprises one or more modifications in addition to those outlined above. Thus, it may be advantageous that one or more proline residues present in the part of the α-amylase variant which is modified is/are replaced with a non-proline residue which may be any of the possible, naturally occurring non-proline residues, and which preferably is an alanine, glycine, serine, threonine, valine or leucine.

Analogously, it may be preferred that one or more cysteine residues present among the amino acid residues with which the parent α-amylase is modified is/are replaced with a non-cysteine residue such as serine, alanine, threonine, glycine, valine or leucine.

Furthermore, a variant of the invention may - either as the only modification or in combination with any of the above outlined modifications - be modified so that one or more Asp and/or Glu present in an amino acid fragment corresponding to the amino acid fragment 185-209 of SEQ ID NO: 4 is replaced by an Asn and/or Gln, respectively. Also of interest is the replacement, in the Termamyl-like α-amylase, of one or more of the Lys residues present in an amino acid fragment corresponding to the amino acid fragment 185-209 of SEQ ID NO: 4 by an Arg.

It will be understood that the present invention encompasses variants incorporating two or more of the above outlined modifications.

Furthermore, it may be advantageous to introduce point-mutations in any of the variants described herein.

Cloning a DNA sequence encoding an α-amylase of the invention

The DNA sequence encoding a parent α-amylase may be isolated from any cell or microorganism producing the α-amylase in question, using various methods well known in the art. First, a genomic DNA and/or cDNA library should be constructed using chromosomal DNA or messenger RNA from the organism that produces the α-amylase to be studied. Then, if the amino acid sequence of the α-amylase is known, homologous, labelled oligonucleotide probes may be synthesized and used to identify α-amylase-encoding clones from a genomic library prepared from the organism in question. Alternatively, a labelled oligonucleotide probe containing sequences homologous to a known α-amylase gene could be used as a probe to identify α-amylase-encoding clones, using hybridization and washing conditions of lower stringency.

Yet another method for identifying α-amylase-encoding clones would involve inserting fragments of genomic DNA into an expression vector, such as a plasmid, transforming α-amylase-negative bacteria with the resulting genomic DNA library, and then plating the transformed bacteria onto agar containing a substrate for α-amylase, thereby allowing clones expressing the α-amylase to be identified.

Alternatively, the DNA sequence encoding the enzyme may be prepared synthetically by established standard methods, e.g. the phosphoroamidite method described by S.L. Beaucage and M.H. Caruthers (1981) or the method described by Matthes et al. (1984). In the phosphoroamidite method, oligonucleotides are synthesized, e.g., in an automatic DNA synthesizer, purified, annealed, ligated and cloned in appropriate vectors.

Finally, the DNA sequence may be of mixed genomic and synthetic origin, mixed synthetic and cDNA origin or mixed genomic and cDNA origin, prepared by ligating fragments of synthetic, genomic or cDNA origin (as appropriate, the fragments corresponding to various parts of the entire DNA sequence), in accordance with standard techniques. The DNA sequence may also be prepared by polymerase chain reaction (PCR) using specific

primers, for instance as described in US 4,683,202 or R.K. Saiki et al. (1988).

Site-directed mutagenesis

Once an a-amylase-encoding DNA sequence has been isolated, and desirable sites for mutation identified, mutations may be introduced using synthetic oligonucleotides. These oligonucleotides contain nucleotide sequences flanking the desired mutation sites; mutant nucleotides are inserted during oligonucleotide synthesis. In a specific method, a single-stranded gap of DNA, bridging the a-amylase-encoding sequence, is created in a vector carrying the a-amylase gene. Then the synthetic nucleotide, bearing the desired mutation, is annealed to a homologous portion of the single-stranded DNA. The remaining gap is then filled in with DNA polymerase I (Klenow fragment) and the construct is ligated using T4 ligase. A specific example of this method is described in Morinaga et al. (1984). US 4,760,025 discloses the introduction of oligonucleotides encoding multiple mutations by performing minor alterations of the cassette. However, an even greater variety of mutations can be introduced at any one time by the Morinaga method, because a multitude of oligonucleotides, of various lengths, can be introduced.

Another method for introducing mutations into a-amylase-encoding DNA sequences is described in Nelson and Long (1989). It involves the 3-step generation of a PCR fragment containing the desired mutation introduced by using a chemically synthesized DNA strand as one of the primers in the PCR reactions. From the PCR-generated fragment, a DNA fragment carrying the mutation may be isolated by cleavage with restriction endonucleases and reinserted into an expression plasmid.

Random Mutagenesis

Random mutagenesis is suitably performed either as localised or region-specific random mutagenesis in at least three parts of the gene translating to the amino acid sequence shown in question, or within the whole gene.

The random mutagenesis of a DNA sequence encoding a parent α-amylase may be conveniently performed by use of any method known in the art.

In relation to the above, a further aspect of the present invention relates to a method for generating a variant of a parent α-amylase, e.g., wherein the variant exhibits altered or increased thermal stability relative to the parent, the method comprising:

(a) subjecting a DNA sequence encoding the parent α-amylase to random mutagenesis,

(b) expressing the mutated DNA sequence obtained in step (a) in a host cell, and

(c) screening for host cells expressing an α-amylase variant which has an altered property (i.e. thermal stability) relative to the parent α-amylase.

Step (a) of the above method of the invention is preferably performed using doped primers.

For instance, the random mutagenesis may be performed by use of a suitable physical or chemical mutagenizing agent, by use of a suitable oligonucleotide, or by subjecting the DNA sequence to PCR generated mutagenesis. Furthermore, the random mutagenesis may be performed by use of any combination of these mutagenizing agents. The mutagenizing agent may, *e.g.*, be one which induces transitions, transversions, inversions, scrambling, deletions, and/or insertions.

Examples of a physical or chemical mutagenizing agent suitable for the present purpose include ultraviolet (UV) irradiation, hydroxylamine, N-methyl-N'-nitro-N-nitrosoguanidine (MNNG), O-methyl hydroxylamine, nitrous acid, ethyl methane sulphonate (EMS), sodium bisulphite, formic acid, and nucleotide analogues. When such agents are used, the mutagenesis is typically performed by incubating the DNA sequence encoding the parent enzyme to be mutagenized in the presence of the mutagenizing agent of choice under suitable conditions for the mutagenesis to take place, and selecting for mutated DNA having the desired properties.

When the mutagenesis is performed by the use of an oligonucleotide, the oligonucleotide may be doped or spiked with the three non-parent nucleotides during the synthesis of the oligonucleotide at the positions which are to be changed. The doping or spiking may be done so that codons for unwanted amino acids are avoided. The doped or spiked oligonucleotide can be incorporated into the DNA encoding the alpha-amylase enzyme by any published technique, using *e.g.* PCR, LCR or any DNA polymerase and ligase as deemed appropriate.

Preferably, the doping is carried out using "constant random doping", in which the percentage of wild-type and mutation in each position is predefined. Furthermore, the doping may be directed toward a preference for the introduction of certain nucleotides, and thereby a preference for the introduction of one or more specific amino acid residues. The doping may be made, *e.g.*, so as to allow for the introduction of 90% wild type and 10% mutations in each position. An additional consideration in the choice of a doping scheme is based on genetic as well as protein-structural constraints. The doping scheme may be made by using the DOPE program which, *inter alia*, ensures that introduction of stop codons is avoided.

When PCR-generated mutagenesis is used, either a chemically treated or non-treated gene encoding a parent α-amylase is subjected to PCR under conditions that increase the misincorporation of nucleotides (Deshler 1992; Leung et al., Technique, Vol.1, 1989, pp. 11-15).

A mutator strain of *E. coli* (Fowler et al., Molec. Gen. Genet., 133, 1974, pp. 179-191), *S. cereviseae* or any other microbial organism may be used for the random mutagenesis of the DNA encoding the α-amylase by, *e.g.*, transforming a plasmid containing the parent glycosylase into the mutator strain, growing the mutator strain with the plasmid and isolating the mutated plasmid from the mutator strain. The mutated plasmid may be subsequently transformed into the expression organism.

The DNA sequence to be mutagenized may be conveniently present in a genomic or cDNA library prepared from an organism expressing the parent alpha-amylase. Alternatively, the DNA sequence may be present on a suitable vector such as a plasmid or a bacteriophage, which as such may be incubated with or other-wise exposed to the mutagenising agent. The DNA to be mutagenized may also be present in a host cell either by being integrated in the genome of said cell or by being present on a vector harboured in the cell. Finally, the DNA to be mutagenized may be in isolated form. It will be understood that the DNA sequence to be subjected to random mutagenesis is preferably a cDNA or a genomic DNA sequence.

In some cases it may be convenient to amplify the mutated DNA sequence prior to performing the expression step b) or the screening step c). Such amplification may be performed in accordance with methods known in the art, the presently preferred method being PCR-generated amplification using oligonucleotide primers prepared on the basis of the DNA or amino acid sequence of the parent enzyme.

Subsequent to the incubation with or exposure to the mutagenising agent, the mutated DNA is expressed by culturing a suitable host cell carrying the DNA sequence under conditions allowing expression to take place. The host cell used for this purpose may be one which has been transformed with the mutated DNA sequence, optionally present on a vector, or one which was carried the DNA sequence encoding the parent enzyme during the mutagenesis treatment. Examples of suitable host cells are the following: gram positive bacteria such as *Bacillus subtilis*, *Bacillus licheniformis*, *Bacillus lentus*, *Bacillus brevis*, *Bacillus stearothermophilus*, *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus coagulans*, *Bacillus circulans*, *Bacillus lautus*, *Bacillus megaterium*, *Bacillus thuringiensis*, *Streptomyces lividans* or *Streptomyces murinus*; and gram-negative bacteria such as *E. coli*.

The mutated DNA sequence may further comprise a DNA sequence encoding functions permitting expression of the mutated DNA sequence.

Localized random mutagenesis

The random mutagenesis may be advantageously localized to a part of the parent α-amylase in question. This may, *e.g.*, be advantageous when certain regions of the enzyme have been identified to be of particular importance for a given property of the enzyme, and when modified are expected to result in a variant having improved properties. Such regions may normally be identified when the tertiary structure of the parent enzyme has been elucidated and related to the function of the enzyme.

The localized, or region-specific, random mutagenesis is conveniently performed by use of PCR generated mutagenesis techniques as described above or any other suitable technique known in the art. Alternatively, the DNA sequence encoding the part of the DNA sequence to be modified may be isolated, *e.g.*, by insertion into a suitable vector, and said part may be subsequently subjected to mutagenesis by use of any of the mutagenesis methods discussed above.

Alternative methods of providing α-amylase variants

Alternative methods for providing variants of the invention include gene shuffling method known in the art including the methods, *e.g.*, described in WO 95/22625 (from Affymax Technologies N.V.) and WO 96/00343 (from Novo Nordisk A/S).

Expression of α-amylase variants of the invention

According to the invention, a DNA sequence encoding the variant produced by methods described above, or by any alternative methods known in the art, can be expressed, in enzyme form, using an expression vector which typically includes control sequences encoding a promoter, operator, ribosome binding site, translation initiation signal, and, optionally, a repressor gene or various activator genes.

The recombinant expression vector carrying the DNA sequence encoding an α-amylase variant of the invention may be any vector which may conveniently be subjected to recombinant DNA procedures, and the choice of vector will often depend on the host cell into which it is to be introduced. Thus, the vector may be an autonomously replicating vector, i.e., a vector which exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, e.g., a plasmid, a bacteriophage or an extrachromosomal element, minichromosome or an artificial chromosome. Alternatively, the vector may be one which, when introduced into a host cell, is integrated into the host cell genome and replicated together with the chromosome(s) into which it has been integrated.

In the vector, the DNA sequence should be operably connected to a suitable promoter sequence. The promoter may be any DNA sequence which shows transcriptional activity in the host cell of choice and may be derived from genes encoding proteins either homologous or heterologous to the host cell. Examples of suitable promoters for directing the transcription of the DNA sequence encoding an α-amylase variant of the invention, especially in a bacterial host, are the promoter of the *lac* operon of *E.coli*, the *Streptomyces coelicolor* agarase gene *dag*A

promoters, the promoters of the *Bacillus licheniformis* α-amylase gene (*amyL*), the promoters of the *Bacillus stearothermophilus* maltogenic amylase gene (*amyM*), the promoters of the *Bacillus amyloliquefaciens* α-amylase (*amyQ*), the promoters of the *Bacillus subtilis* xylA and xylB genes etc. For transcription in a fungal host, examples of useful promoters are those derived from the gene encoding *A. oryzae* TAKA amylase, *Rhizomucor miehei* aspartic proteinase, *A. niger* neutral α-amylase, *A. niger* acid stable α-amylase, *A. niger* glucoamylase, *Rhizomucor miehei* lipase, *A. oryzae* alkaline protease, *A. oryzae* triose phosphate isomerase or *A. nidulans* acetamidase.

The expression vector of the invention may also comprise a suitable transcription terminator and, in eukaryotes, polyadenylation sequences operably connected to the DNA sequence encoding the α-amylase variant of the invention. Termination and polyadenylation sequences may suitably be derived from the same sources as the promoter.

The vector may further comprise a DNA sequence enabling the vector to replicate in the host cell in question. Examples of such sequences are the origins of replication of plasmids pUC19, pACYC177, pUB110, pE194, pAMB1 and pIJ702.

The vector may also comprise a selectable marker, e.g. a gene the product of which complements a defect in the host cell, such as the *dal* genes from *B. subtilis* or *B. licheniformis*, or one which confers antibiotic resistance such as ampicillin, kanamycin, chloramphenicol or tetracyclin resistance. Furthermore, the vector may comprise *Aspergillus* selection markers such as amdS, argB, niaD and sC, a marker giving rise to hygromycin resistance, or the selection may be accomplished by co-transformation, e.g. as described in WO 91/17243.

While intracellular expression may be advantageous in some respects, e.g. when using certain bacteria as host cells, it is generally preferred that the expression is extracellular. In general, the *Bacillus* α-amylases mentioned herein comprise a preregion permitting secretion of the expressed protease into the culture medium. If desirable, this preregion may be replaced

by a different preregion or signal sequence, conveniently accomplished by substitution of the DNA sequences encoding the respective preregions.

The procedures used to ligate the DNA construct of the invention encoding an α-amylase variant, the promoter, terminator and other elements, respectively, and to insert them into suitable vectors containing the information necessary for replication, are well known to persons skilled in the art (cf., for instance, Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor, 1989).

The cell of the invention, either comprising a DNA construct or an expression vector of the invention as defined above, is advantageously used as a host cell in the recombinant production of an α-amylase variant of the invention. The cell may be transformed with the DNA construct of the invention encoding the variant, conveniently by integrating the DNA construct (in one or more copies) in the host chromosome. This integration is generally considered to be an advantage as the DNA sequence is more likely to be stably maintained in the cell. Integration of the DNA constructs into the host chromosome may be performed according to conventional methods, *e.g.* by homologous or heterologous recombination. Alternatively, the cell may be transformed with an expression vector as described above in connection with the different types of host cells.

The cell of the invention may be a cell of a higher organism such as a mammal or an insect, but is preferably a microbial cell, *e.g.*, a bacterial or a fungal (including yeast) cell.

Examples of suitable bacteria are grampositive bacteria such as *Bacillus subtilis*, *Bacillus licheniformis*, *Bacillus lentus*, *Bacillus brevis*, *Bacillus stearothermophilus*, *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus coagulans*, *Bacillus circulans*, *Bacillus lautus*, *Bacillus megaterium*, *Bacillus thuringiensis*, or *Streptomyces lividans* or *Streptomyces murinus*, or gramnegative bacteria such as *E.coli*. The transformation of the bacteria may, for instance, be effected by protoplast transformation or by using competent cells in a manner known *per se*.

The yeast organism may favourably be selected from a species of *Saccharomyces* or *Schizosaccharomyces*, *e.g.* *Saccharomyces cerevisiae*. The filamentous fungus may advantageously belong to a species of *Aspergillus*, *e.g.* *Aspergillus oryzae* or *Aspergillus niger*. Fungal cells may be transformed by a process involving protoplast formation and transformation of the protoplasts followed by regeneration of the cell wall in a manner known *per se*. A suitable procedure for transformation of *Aspergillus* host cells is described in EP 238 023.

In a yet further aspect, the present invention relates to a method of producing an α-amylase variant of the invention, which method comprises cultivating a host cell as described above under conditions conducive to the production of the variant and recovering the variant from the cells and/or culture medium.

The medium used to cultivate the cells may be any conventional medium suitable for growing the host cell in question and obtaining expression of the α-amylase variant of the invention. Suitable media are available from commercial suppliers or may be prepared according to published recipes (*e.g.*, as described in catalogues of the American Type Culture Collection).

The α-amylase variant secreted from the host cells may conveniently be recovered from the culture medium by well-known procedures, including separating the cells from the medium by centrifugation or filtration, and precipitating proteinaceous components of the medium by means of a salt such as ammonium sulphate, followed by the use of chromatographic procedures such as ion exchange chromatography, affinity chromatography, or the like.

Industrial Applications

The α-amylase variants of this invention possesses valuable properties allowing for a variety of industrial applications. An enzyme variant of the invention are applicable as a component in washing, dishwashing and hard-surface cleaning detergent compositions. Numerous variants are particularly useful in the production of sweeteners and ethanol from starch, and/or for

textile desizing. Conditions for conventional starch- conversion processes, including starch liquefaction and/or saccharification processes, are described in, *e.g.*, US 3,912,590 and in EP patent publications Nos. 252,730 and 63,909.

<u>Production of sweeteners from starch</u>:

A "traditional" process for conversion of starch to fructose syrups normally consists of three consecutive enzymatic processes, viz. a liquefaction process followed by a saccharification process and an isomerization process. During the liquefaction process, starch is degraded to dextrins by an α-amylase (*e.g.* Termamyl™) at pH values between 5.5 and 6.2 and at temperatures of 95-160°C for a period of approx. 2 hours. In order to ensure an optimal enzyme stability under these conditions, 1 mM of calcium is added (40 ppm free calcium ions).

After the liquefaction process the dextrins are converted into dextrose by addition of a glucoamylase (*e.g.* AMG™) and a debranching enzyme, such as an isoamylase or a pullulanase (*e.g.* Promozyme™). Before this step the pH is reduced to a value below 4.5, maintaining the high temperature (above 95°C), and the liquefying α-amylase activity is denatured. The temperature is lowered to 60°C, and glucoamylase and debranching enzyme are added. The saccharification process proceeds for 24-72 hours.

After the saccharification process the pH is increased to a value in the range of 6-8, preferably pH 7.5, and the calcium is removed by ion exchange. The dextrose syrup is then converted into high fructose syrup using, *e.g.*, an immmobilized glucoseisomerase (such as Sweetzyme™).

At least 1 enzymatic improvements of this process could be envisaged. Reduction of the calcium dependency of the liquefying α-amylase. Addition of free calcium is required to ensure adequately high stability of the α-amylase, but free calcium strongly inhibits the activity of the glucoseisomerase and needs to be removed, by means of an expensive unit operation, to an extent which reduces the level of free calcium to below 3-5 ppm. Cost savings could be obtained if such an

operation could be avoided and the liquefaction process could be performed without addition of free calcium ions.

To achieve that, a less calcium-dependent Termamyl-like α-amylase which is stable and highly active at low concentrations of free calcium (< 40 ppm) is required. Such a Termamyl-like α-amylase should have a pH optimum at a pH in the range of 4.5-6.5, preferably in the range of 4.5-5.5.

Detergent compositions

As mentioned above, variants of the invention may suitably be incorporated in detergent compositions. Reference is made, for example, to WO 96/23874 and WO 97/07202 for further details concerning relevant ingredients of detergent compositions (such as laundry or dishwashing detergents), appropriate methods of formulating the variants in such detergent compositions, and for examples of relevant types of detergent compositions.

Detergent compositions comprising a variant of the invention may additionally comprise one or more other enzymes, such as a lipase, cutinase, protease, cellulase, peroxidase or laccase, and/or another α-amylase.

α-amylase variants of the invention may be incorporated in detergents at conventionally employed concentrations. It is at present contemplated that a variant of the invention may be incorporated in an amount corresponding to 0.00001-1 mg (calculated as pure, active enzyme protein) of α-amylase per liter of wash/dishwash liquor using conventional dosing levels of detergent.

MATERIALS AND METHODS

Enzymes:

LE174 hybrid alpha-amylase variant: LE174 is a hybrid Termamyl-like alpha-amylase being identical to the Termamyl sequence, i.e., the *Bacillus licheniformis* α-amylase shown in SEQ ID NO: 4, except that the N-terminal 35 amino acid residues (of the mature protein) has been replaced by the N-terminal 33 residues of BAN (mature protein), i.e., the *Bacillus amyloliquefaciens*

alpha-amylase shown in SEQ ID NO: 5, which further havefollowing mutations: H156Y+A181T+N190F+A209V+Q264S (using the numbering in SEQ ID NO: 4).

Construction of pSNK101

This *E. coli/Bacillus* shuttle vector can be used to introduce mutations without expression of α-amylase in *E. coli* and then be modified in such way that the α-amylase is active in Bacillus. The vector was constructed as follows: The α-amylase gene in the pX vector (pDN1528 with the following alterations within amyL: BAN(1-33), H156Y, A181T, N190F, A209V, Q264S, the plasmid pDN1528 is further described in Example 1) was inactivated by interruption in the PstI site in the 5´coding region of the alpha-amylase gene by a 1.2 kb fragment containing an *E. coli* origin fragment. This fragment was amplified from the pUC19 (GenBank Accession #:X02514) using the forward primer 1: 5´-gacctgcagtcaggcaacta-3´ (SEQ ID NO: 28) and the reverse primer 1: 5´-tagagtcgacctgcaggcat-3´ (SEQ ID NO: 29). The PCR amplicon and the pX plasmid containing the α-amylase gene were digested with PstI at 37°C for 2 hours. The pX vector fragment and the *E. coli* origin amplicon were ligated at room temperature. for 1 hour and transformed in *E. coli* by electrotransformation. The resulting vector is designated pSnK101.

This *E. coli/Bacillus* shuttle vector can be used to introduce mutations without expression of α-amylase in *E. coli* and then be modified in such way that the α-amylase is active in Bacillus. The vector was constructed as follows: The α-amylase gene in the pX vector (pDN1528 with the following alterations within amyL: BAN(1-33), H156Y+A181T+N190F+A209V+Q264S, the plasmid pDN1528 is further described in Example 1) was inactivated by interruption in the PstI site in the 5´coding region of the alpha-amylase gene by a 1.2 kb fragment containing an *E. coli* origin fragment. This fragment was amplified from the pUC19 (GenBank Accession

#:X02514) using the forward primer 2: 5´-gacctgcagtcaggcaacta-3´ (SEQ ID NO: 30) and the reverse primer 2: 5´-tagagtcgacctgcaggcat-3´ (SEQ ID NO: 31). The PCR amplicon and the pX plasmid containing the α-amylase gene were digested with PstI at 37°C for 2 hours. The pX vector fragment and the *E. coli* origin amplicon were ligated at room temperature. for 1 hour and transformed in *E. coli* by electrotransformation. The resulting vector is designated pSnK101.

Low pH filter assay

Bacillus libraries are plated on a sandwich of cellulose acetate (OE 67, Schleicher & Schuell, Dassel, Germany) - and nitrocellulose filters (Protran-Ba 85, Schleicher & Schuell, Dassel, Germany) on TY agar plates with 10 µg/ml chloramphenicol at 37°C for at least 21 hrs. The cellulose acetate layer is located on the TY agar plate.

Each filter sandwich is specifically marked with a needle after plating, but before incubation in order to be able to localize positive variants on the filter and the nitrocellulose filter with bound variants is transferred to a container with citrate buffer, pH 4.5 and incubated at 90°C for 15 min. The cellulose acetate filters with colonies are stored on the TY-plates at room temperature until use. After incubation, residual activity is detected on assay plates containing 1% agarose, 0.2% starch in citrate buffer, pH 6.0. The assay plates with nitrocellulose filters are marked the same way as the filter sandwich and incubated for 2 hours. at 50°C. After removal of the filters the assay plates are stained with 10% Lugol solution. Starch degrading variants are detected as white spots on dark blue background and then identified on the storage plates. Positive variants are rescreened twice under the same conditions as the first screen.

Secondary screening

Positive transformants after rescreening are picked from the storage plate and tested in a secondary plate assay.

Positive transformants are grown for 22 hours at 37°C in 5 ml LB + chloramphenicol. The Bacillus culture of each positive transformant and a control LE174 variant were incubated in citrate buffer, pH 4.5 at 90°C and samples were taken at 0,10,20,30,40,60 and 80 minutes. A 3 microliter sample was spotted on a assay plate. The assay plate was stained with 10% Lugol solution. Improved variants were seen as variants with higher residual activity detected as halos on the assay plate than the backbone. The improved variants are determined by nucleotide sequencing.

Fermentation and purification of α-amylase variants

A *B. subtilis* strain harbouring the relevant expression plasmid is streaked on a LB-agar plate with 15 μg/ml chloramphenicol from -80°C stock, and grown overnight at 37°C. The colonies are transferred to 100 ml BPX media supplemented with 15 μg/ml chloramphenicol in a 500 ml shaking flask.

Composition of BPX medium:

Potato starch	100 g/l
Barley flour	50 g/l
BAN 5000 SKB	0.1 g/l
Sodium caseinate	10 g/l
Soy Bean Meal	20 g/l
Na_2HPO_4, 12 H_2O	9 g/l
Pluronic™	0.1 g/l

The culture is shaken at 37°C at 270 rpm for 5 days.

Cells and cell debris are removed from the fermentation broth by centrifugation at 4500 rpm in 20-25 minutes. Afterwards the supernatant is filtered to obtain a completely clear solution. The filtrate is concentrated and washed on a UF-filter (10000 cut off membrane) and the buffer is changed to 20mM Acetate pH 5.5. The UF-filtrate is applied on a S-sepharose F.F. and elution is carried out by step elution with 0.2M NaCl in the same buffer. The eluate is dialysed against 10mM Tris, pH 9.0 and applied on a Q-sepharose F.F. and eluted with a linear gradient from 0-0.3M NaCl over 6 column volumes. The fractions

which contain the activity (measured by the Phadebas assay) are pooled, pH was adjusted to pH 7.5 and remaining color was removed by a treatment with 0.5% W/vol. active coal in 5 minutes.

Stability determination

All the stability trials are made using the same set up. The method is:

The enzyme is incubated under the relevant conditions (1-4). Samples are taken at 0, 5, 10, 15 and 30 minutes and diluted 25 times (same dilution for all taken samples) in assay buffer (0.1M 50mM Britton buffer pH 7.3) and the activity is measured using the Phadebas assay (Pharmacia) under standard conditions pH 7.3, 37°C.

The activity measured before incubation (0 minutes) is used as reference (100%). The decline in percent is calculated as a function of the incubation time. The table shows the residual activity after 30 minutes of incubation.

Activity determination - (KNU)

One Kilo alpah-amylase Unit (1 KNU) is the amount of enzyme which breaks down 5.26 g starch (Merck, Amylum Solubile, Erg. B 6, Batch 9947275) per hour in Novo Nordisk's standard method for determination of alpha-amylase based upon the following condition:

Substrate	soluble starch
Calcium content in solvent	0.0043 M
Reaction time	7-20 minutes
Temperature	37°C
pH	5.6

Detailed description of Novo Nordisk's analytical method (AF 9) is available on request.

Specific activity determination

<u>Assay for α-Amylase Activity</u>

α-amylase activity is determined by a method employing Phadebas® tablets as substrate. Phadebas tablets

(Phadebas® Amylase Test, supplied by Pharmacia Diagnostic) contain a cross-linked insoluble blue-coloured starch polymer which has been mixed with bovine serum albumin and a buffer substance and tabletted.

For every single measurement one tablet is suspended in a tube containing 5 ml 50 mM Britton-Robinson buffer (50 mM acetic acid, 50 mM phosphoric acid, 50 mM boric acid, 0.1 mM $CaCl_2$, pH adjusted to the value of interest with NaOH). The test is performed in a water bath at the temperature of interest. The α-amylase to be tested is diluted in x ml of 50 mM Britton-Robinson buffer. 1 ml of this α-amylase solution is added to the 5 ml 50 mM Britton-Robinson buffer. The starch is hydrolysed by the α-amylase giving soluble blue fragments. The absorbance of the resulting blue solution, measured spectrophotometrically at 620 nm, is a function of the α-amylase activity.

It is important that the measured 620 nm absorbance after 10 or 15 minutes of incubation (testing time) is in the range of 0.2 to 2.0 absorbance units at 620 nm. In this absorbance range there is linearity between activity and absorbance (Lambert-Beer law). The dilution of the enzyme must therefore be adjusted to fit this criterion. Under a specified set of conditions (temp., pH, reaction time, buffer conditions) 1 mg of a given α-amylase will hydrolyse a certain amount of substrate and a blue colour will be produced. The colour intensity is measured at 620 nm. The measured absorbance is directly proportional to the specific activity (activity/mg of pure α-amylase protein) of the α-amylase in question under the given set of conditions.

EXAMPLES

Example 1.

<u>Construction, by random mutagenesis, of Termamyl-like LE174 α-amylase variants having an improved stability at low pH and a reduced dependency on calcium ions for stability compared to the parent enzyme.</u>

Random mutagenesis

To improve the stability at low pH and low calcium concentration of the parent LE174 α-amylase variant random mutagenesis in preselected regions was performed.

The regions were:

Region:	Residue:
SERI	A425-Y438
SERII	W411-L424
SERIII	G397-G410
SERV	T369-H382
SERVII	G310-F323
SERIX	L346-P359

For each six region, random oligonucleotides are synthesized using the same mutation rate (97 % backbone and 1% of each of the three remaining nucleotides giving 3% mutations) in each nucleotide position in the above regions, e.g., 1. position in condon for A425: 97%C, 1%A, 1%T, 1%G. The six random oligonucleotides and if used complementary SOE helping primers are shown in tables1-6: with the four distribution of nucleotides below.

Table 1.

RSERI: 5'-GC GTT TTG CCG GCC GAC ATA 312 234 322 243 333 133 444 233 423 242 212 211 243 343 CAA ACC TGA ATT-3' (SEQ ID NO: 15)

Table 2.

RSERII: 5'-GC GTT TTG CCG GCC GAC ATA CAT TCG CTT TGC CCC ACC GGG TCC GTC TGT TAT TAA TGC CGC 311 133 241 122 243 113 341 432 423 433 223 332 242 331 GCC GAC AAT GTC ATG GTG-3' (SEQ ID NO: 16)

Table 3.

RSERIII: 5'-GTC GCC TTC CCT TGT CCA 433 413 112 423 124 424 423 411 121 123 124 324 243 233 GTA CGC ATA CTG TTT TCT-3' (SEQ ID NO: 17)

Helping primer FSERIII: 5'-TGG ACA AGG GAA GGC GAC AG-3' (SEQ ID NO: 18)

Table 4.

RSERV: 5-TAA GAT CGG TTC AAT TTT 424 222 311 443 144 112 223 434 324 441 423 233 222 342 CCC GTA CAT ATC CCC GTA GAA-3 (SEQ ID NO: 19)

Helping primer FSERV: 5-AAA ATT GAA CCG ATC TTA-3 (SEQ ID NO: 20)

Table 5.

FSERVII: 5'-TT CCA TGC TGC ATC GAC ACA GGG AGG CGG CTA TGA TAT GAG GAA ATT GCT GAA 344 213 442 342 223 311 431 233 422 411 123 442 213 122 TGT CGA TAA CCA-3' (SEQ ID NO: 21)

Helping primer RSERVII: 5´- TGT CGA TGC AGC ATG GAA - 3´ (SEQ ID NO: 22)

Table 6.

FSERIX: 5'-GT CCA AAC ATG GTT TAA GCC 432 243 221 343 222 212 232 313 114 441 123 244 121 333 TCA GGT TTT CTA CGG GGA-3' (SEQ ID NO: 23)

Helping primer RSERIX: 5´-GGC TTA AAC CAT GTT TGG AC-3´ (SEQ ID NO: 24)

Distribution of nucleotides in each mutated nucleotide position

1:97%A, 1%T, 1%C, 1%G

2:97%T, 1%A, 1%C, 1%G

3:97%C, 1%A, 1%T, 1%G

4:97%G, 1%A, 1%T, 1%C

Construction of plasmid libraries

Two approximately 1.4 kb fragments were PCR amplified using the primer 1B: 5'-CGA TTG CTG ACG CTG TTA TTT GCG-3' and the random oligonucleotide apparent from table 1, respectively the random oligonucleotide apparent from table 2. The vector pSnK101 and the PCR fragments were digested with EcoRV and EagI for 2 hours. The approximately 3.6 kb vector fragment and the approximately 1.3 kb PCR fragments was purified and ligated overnight and transformed in to *E.coli* and then further transformed into a *Bacillus* host starin as described below. The random oligonucleotides apparent from Tables 3-6 (which by a common term is designated aSER and bSER in Fig. 2) for each region and specific *B. licheniformis* primers 1B (SEQ ID NO: 26) and #63: 5'-CTA TCT TTG AAC ATA AAT TGA AAC C-3' (SEQ ID NO: 27) covering the EcoRV and the EagI sites in the LE174 sequence are used to generate PCR-library-fragments by the overlap extension method (Horton et al., Gene, 77 (1989), pp. 61-68) Figure 2 shows the PCR strategy. The PCR fragments are cloned in the *E. coli/Bacillus* shuttle vector pSNK101 (see Materials and Methods) enabling mutagenesis in *E. coli* and immediate expression in *Bacillus subtilis* preventing lethal accumulation of amylases in *E. coli*. After establishing the cloned PCR fragments in *E. coli*, a modified pUC19 fragment is digested out of the plasmid and the promoter and the mutated Termamyl gene is physically connected and expression can take place in the *Bacillus* host .

Screening

The six libraries were screened in the low pH filter assays described in the "Material and Methods" section above.

All variants listed in the table in Example 2 below was prepared as described in Example 1.

EXAMPLE 2

Measurement of stability

Normally, industrial liquefaction processes is run at pH 6.0-6.2 with addition of about 40 ppm free calcium in order to

improve the stability at 95°C-105°C. Variants of the invention have been made in order to improve the stability at

1. lower pH than pH 6.2 and/or
2. at free calcium levels lower than 40ppm free calcium.

An assay which measures the stability at acidic pH, pH 5.0, in the presence of 5ppm free calcium was used to measure the increase in stability.

10 μg of the variant was incubated under the following conditions: A 0.1 M acetate solution, pH adjusted to pH 5.0, containing 5ppm calcium and 5% w/w common corn starch (free of calcium). Incubation was made in a water bath at 95°C for 30 minutes.

Results:

Increased stability at pH 5.0, 5 ppm calcium incubated at 95°C

MINUTES OF INCUBATION	LE174 WITH K176R+ I201F+ H205N	LE174 WITH K176R+ I201F+ H205N+ E376K+ A420R	LE174 WITH K176R+ I201F+ H205N+ S417T+ A420Q	LE174 WITH K176R+ I201F+ H205N+ S356A+ Y358F
0	100	100	100	100
5	65	61	66	66
10	58	53	60	59
15	51	48	55	56
30	36	39	45	49

Specific activity determination.

The specific activity was determined using the Phadebas assay (Pharmacia) (described above) as activity/mg enzyme. The activity was determined using the α-amylase assay described in the Materials and Methods section herein.

LE174 with the following substitutions:

K176R+I201F+H205N
Specific activity determined: 13400NU/mg

LE174 with the following substitutions:
K176R+I201F+H205N+E376K+A420R:
Specific activity determined: 14770NU/mg

LE174 with the following substitutions:
K176R+I201F+H205N+S417T+A420Q:
Specific activity determined:16670NU/mg

LE174 with the following substitutions:
K176R+I201F+H205N+S356A+Y358F:
Specific activity determined:15300NU/mg

REFERENCES CITED

Klein, C., et al., *Biochemistry 1992*, **31**, 8740-8746,

Mizuno, H., et al., *J. Mol. Biol.* (1993) **234**, 1282-1283,

Chang, C., et al, *J. Mol. Biol.* (1993) **229**, 235-238,

Larson, S.B., *J. Mol. Biol.* (1994) **235**, 1560-1584,

Lawson, C.L., *J. Mol. Biol.* (1994) **236**, 590-600,

Qian, M., et al., *J. Mol. Biol.* (1993) **231**, 785-799,

Brady, R.L., et al., *Acta Crystallogr. sect. B*, **47**, 527-535,

Swift, H.J., et al., *Acta Crystallogr. sect. B*, **47**, 535-544

A. Kadziola, Ph.D. Thesis: "An alpha-amylase from Barley and its Complex with a Substrate Analogue Inhibitor Studied by X-ray Crystallography", Department of Chemistry University of Copenhagen 1993

MacGregor, E.A., Food Hydrocolloids, 1987, Vol.1, No. 5-6, p.

B. Diderichsen and L. Christiansen, Cloning of a maltogenic α-amylase from *Bacillus stearothermophilus*, FEMS Microbiol. letters: 56: pp. 53-60 (1988)

Hudson et al., Practical Immunology, Third edition (1989), Blackwell Scientific Publications,

Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor, 1989

S.L. Beaucage and M.H. Caruthers, Tetrahedron Letters 22, 1981, pp. 1859-1869

Matthes et al., The EMBO J. 3, 1984, pp. 801-805.

R.K. Saiki et al., Science 239, 1988, pp. 487-491.

Morinaga et al., (1984, Biotechnology 2:646-639)

Nelson and Long, Analytical Biochemistry 180, 1989, pp. 147-151

Hunkapiller et al., 1984, Nature 310:105-111

R. Higuchi, B. Krummel, and R.K. Saiki (1988). A general method of *in vitro* preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. *Nucl. Acids Res.* **16**:7351-7367.

Dubnau et al., 1971, J. Mol. Biol. 56, pp. 209-221.

Gryczan et al., 1978, J. Bacteriol. 134, pp. 318-329.

S.D. Erlich, 1977, Proc. Natl. Acad. Sci. 74, pp. 1680-1682.

Boel et al., 1990, Biochemistry 29, pp. 6244-6249.

CLAIMS

1. A variant of a parent Termamyl-like α-amylase, which variant α-amylase has been altered in comparison to the parent α-amylase in one or more solvent exposed amino acid residues on the surface of the α-amylase to increase the overall hydrophobicity of the α-amylase and/or to increase the overall numbers of methyl groups in the sidechains of said solvent exposed amino acid residues on the surface.

2. The variant according to claim 1, wherein one or more solvent exposed amino acid residues on a concav surface with inwards bend are altered to more hydrophobic amino acid residues.

3. The variant according to claim 1, wherein one or more solvent exposed amino acid residues on a convex surface are altered to increase the number of methyl groups in the sidechain.

4. A variant of a parent Termamyl-like α-amylase, comprising an alteration at one or more positions selected from the group of:
E376, S417, A420, S356, Y358;
wherein (a) the alteration(s) are independently
(i) an insertion of an amino acid downstream of the amino acid which occupies the position,
(ii) a deletion of the amino acid which occupies the position, or
(iii) a substitution of the amino acid which occupies the position with a different amino acid,
(b) the variant has α-amylase activity and (c) each position corresponds to a position of the amino acid sequence of the parent Termamyl-like α-amylase having the amino acid sequence of SEQ ID NO: 4.

5. The variant according to claim 4, which variant has an alteration in one or more solvent exposed amino acid residues as defined in any of claims 1-3.

6. The variant of any of claims 1-5, wherein the parent Termamyl-like α-amylase is derived from a strain of *B. licheniformis, B. amyloliquefaciens, B. stearothermophilus, Bacillus* sp. NCIB 12289, NCIB 12512, NCIB 12513 or DSM 9375.

7. The variant accordin to claim 6, wherein the parent α-amylase is derived from *B. licheniformis* strain ATCC 27811.

8. The variant according to claims 1-6, wherein the parent Termamyl-like α-amylase is any of the α-amylases selected from the group depicted in SEQ ID NOS: 1, 2, 3, 4, 5, 6, 7, and 8.

9. The variant according to any of claims 1-8, wherein the parent Termamyl-like α-amylase has an amino acid sequence which has a degree of identity to SEQ ID NO: 4 of at least 65%, preferably 70%, more preferably at least 80%, even more preferably at least about 90%, even more preferably at least 95%, even more preferably at least 97%, and even more preferably at least 97%.

10. The variant according to any of claims 1-10, wherein the parent Termamyl-like α-amylase is encoded by a nucleic acid sequence which hydridizes under medium, preferred high stringency conditions, with the nucleic acid sequence of SEQ ID NO: 12.

11. The variant accordint to claims 1-10, wherein the parent Termamyl-like α-amylase is a hybrid of the *B. licheniformis* α-amylase shown in SEQ ID NO: 4 and *B. amyloliquefaciens* α-amylase shown in SEQ ID NO: 5.

12. The variant according to claim 11, wherein the parent hybrid Termayl-like α-amylase is LE174.

13. The variant according to any of claims 1-12, wherein the parent α-amylase further has a mutation in one or more of the following positions: K176, I201 and H205 (using the numbering in SEQ ID NO: 4).

14. The variant according to claim 13, wherein the parent α-amylase has one or more the following substitutions: K176R, I201F and/or H205N (using the numbering in SEQ ID NO: 4).

15. The variant according to claim 14, wherein the parent α-amylase has the following substitutions: K176R+I201F+H205N (using the numbering in SEQ ID NO: 4).

16. The variant according to claims 1 to 15, wherein the variant has increased stability at pHs below 7.0 (acidic pH) and/or at low calcium concentration and/or at temperatures in the range from 95 to 160°C (high temperatures) relative to the parent α-amylase.

17. The variant according to any of claims 1 to 16, which variant has one or more of the following substitutions: E376K,S417T,A420Q,R,S356A,Y358F.

18. A DNA construct comprising a DNA sequence encoding an α-amylase variant according to any one of claims 1 to 17.

19. A recombinant expression vector which carries a DNA construct according to claim 18.

20. A cell which is transformed with a DNA construct according to claim 18 or a vector according to claim 19.

21. A cell according to claim 20, which is a microorganism.

22. A cell according to claim 21, which is a bacterium or a fungus.

23. The cell according to claim 22, which is a grampositive bacterium such as *Bacillus subtilis*, *Bacillus licheniformis*, *Bacillus lentus*, *Bacillus brevis*, *Bacillus stearothermophilus*, *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus coagulans*, *Bacillus circulans*, *Bacillus lautus* or *Bacillus thuringiensis*.

24. A detergent additive comprising an α-amylase variant according to any one of claims 1 to 17, optionally in the form of a non-dusting granulate, stabilised liquid or protected enzyme.

25. A detergent additive according to claim 24 which contains 0.02-200 mg of enzyme protein/g of the additive.

26. A detergent additive according to claims 24 or 25, which additionally comprises another enzyme such as a protease, a lipase, a peroxidase, another amylolytic enzyme and/or a cellulase.

27. A detergent composition comprising an α-amylase variant according to any of claims 1 to 17.

28. A detergent composition according to claim 27 which additionally comprises another enzyme such as a protease, a lipase, a peroxidase, another amylolytic enzyme and/or a cellulase.

29. A manual or automatic dishwashing detergent composition comprising an α-amylase variant according to any of claims 1 to 17.

30. A dishwashing detergent composition according to claim 29 which additionally comprises another enzyme such as a protease, a lipase, a peroxidase, another amylolytic enzyme and/or a cellulase.

31. A manual or automatic laundry washing composition comprising

an α-amylase variant according to any of claims 1 to 17.

32. A laundry washing composition according to claim 31, which additionally comprises another enzyme such as a protease, a lipase, a peroxidase, an amylolytic enzyme and/or a cellulase.

33. A composition comprising:

(i) a mixture of the α-amylase from *B. licheniformis* having the sequence shown in SEQ ID NO: 4 with one or more variants according to any of claims 1 to 17 derived from (as the parent Termamyl-like α-amylase) the *B. stearothermophilus* α-amylase having the sequence shown in SEQ ID NO: 3; or

(ii) a mixture of the α-amylase from *B. stearothermophilus* having the sequence shown in SEQ ID NO: 3 with one or more variants according to any of claims 1 to 17 derived from one or more other parent Termamyl-like α-amylases; or

(iii) a mixture of one or more variants according any of claims 1 to 17 derived from (as the parent Termamyl-like α-amylase) the *B. stearothermophilus* α-amylase having the sequence shown in SEQ ID NO: 3 with one or more variants according to the invention derived from one or more other parent Termamyl-like α-amylases.

34. The composition comprising a variant of any of claims 1 to 17 wherein the parent α-amylase is a hybrid alpha-amylase comprising a N-terminal part of the *B. amyloliquefaciens* α-amylase shown in SEQ ID NO: 5 and a C-terminal part of the *B. licheniformis* α-amylase shown in SEQ ID NO: 4.

35. The composition according to claim 34, wherein the parent hybrid Termamyl-like α-amylase is LE174

36. The composition according to claims 35, wherein the parent Termamyl-like α-amylase is LE174 with an alterantion in one or more of the following positions: K176, I201 and H205.

37. The composition according to claims 36, wherein the parent Termamyl-like α-amylase is LE174 with one or more of the following substitutions: K176R, I201F and H205N.

38. Use of an α-amylase variant according to any one of claims 1 to 17 or a composition according to claims 33 to 37 for washing and/or dishwashing.

39. Use of an α-amylase variant according to any one of claims 1 to 17 or a composition according to claims 33 to 37 for textile desizing.

40. Use of an α-amylase variant according to any of claims 1 to 17 or a composition according to claims 33 to 37 for starch liquefaction.

41. A method for generating a variant of a parent Termamyl-like α-amylase, which variant exhibits increased stability at high temperatures relative to the parent, the method comprising:
(a) subjecting a DNA sequence encoding the parent Termamyl-like α-amylase to random mutagenesis,
(b) expressing the mutated DNA sequence obtained in step (a) in a host cell, and
(c) screening for host cells expressing a mutated α-amylase which has increased stability at high temperatures relative to the parent Termamyl-like α-amylase.

```
    1                                                     50
1   HHNGTNGTMM QYFEWHLPND GNHWNRLRDD ASNLRNRGIT AIWIPPAWKG
2   ..NGTNGTMM QYFEWYLPND GNHWNRLRSD ASNLKDKGIS AVWIPPAWKG
3   HHNGTNGTMM QYFEWYLPND GNHWNRLRDD AANLKSKGIT AVWIPPAWKG
4   ....VNGTLM QYFEWYTPND GQHWKRLQND AEHLSDIGIT AVWIPPAYKG
5   ..ANLNGTLM QYFEWYMPND GQHWRRLQND SAYLAEHGIT AVWIPPAYKG
6   .AAPFNGTMM QYFEWYLPDD GTLWTKVANE ANNLSSLGIT ALWLPPAYKG

    51                                                   100
1   TSQNDVGYGA YDLYDLGEFN QKGTVRTKYG TRSQLESAIH ALKNNGVQVY
2   ASQNDVGYGA YDLYDLGEFN QKGTIRTKYG TRNQLQAAVN ALKSNGIQVY
3   TSQNDVGYGA YDLYDLGEFN QKGTVRTKYG TRNQLQAAVT SLKNNGIQVY
4   LSQSDNGYGP YDLYDLGEFQ QKGTVRTKYG TKSELQDAIG SLHSRNVQVY
5   TSQADVGYGA YDLYDLGEFH QKGTVRTKYG TKGELQSAIK SLHSRDINVY
6   TSRSDVGYGV YDLYDLGEFN QKGTVRTKYG TKAQYLQAIQ AAHAAGMQVY

    101                                                  150
1   GDVVMNHKGG ADATENVLAV EVNPNNRNQE ISGDYTIEAW TKFDFPGRGN
2   GDVVMNHKGG ADATEMVRAV EVNPNNRNQE VSGEYTIEAW TKFDFPGRGN
3   GDVVMNHKGG ADGTEIVNAV EVNRSNRNQE TSGEYAIEAW TKFDFPGRGN
4   GDVVLNHKAG ADATEDVTAV EVNPANRNQE TSEEYQIKAW TDFRFPGRGN
5   GDVVINHKGG ADATEDVTAV EVDPADRNRV ISGEHLIKAW THFHFPGRGS
6   ADVVFDHKGG ADGTEWVDAV EVNPSDRNQE ISGTYQIQAW TKFDFPGRGN

    151                                                  200
1   TYSDFKWRWY HFDGVDWDQS RQFQNRIYKF RGDGKAWDWE VDSENGNYDY
2   THSNFKWRWY HFDGVDWDQS RKLNNRIYKF RGDGKGWDWE VDTENGNYDY
3   NHSSFKWRWY HFDGTDWDQS RQLQNKIYKF RGTGKAWDWE VDTENGNYDY
4   TYSDFKWHWY HFDGADWDES RKI.SRIFKF RGEGKAWDWE VSSENGNYDY
5   TYSDFKWHWY HFDGTDWDES RKL.NRIYKF ..QGKAWDWE VSNENGNYDY
6   TYSSFKWRWY HFDGVDWDES RKL.SRIYKF RGIGKAWDWE VDTENGNYDY
```

Fig. 1

```
     201                                                        250
  1  LMYADVDMDH PEVVNELRRW GEWYTNTLNL DGFRIDAVKH IKYSFTRDWL
  2  LMYADIDMDH PEVVNELRNW GVWYTNTLGL DGFRIDAVKH IKYSFTRDWS
  3  LMYADVDMDH PEVIHELRNW GVWYTNTLNL DGFRIDAVKH IKYSFTRDWL
  4  LMYADVDYDH PDVVAETKKW GIWYANELSL DGFRIDAAKH IKFSFLRDWV
  5  LMYADIDYDH PDVAAEIKRW GTWYANELQL DGFRLDAVKH IKFSFLRDWV
  6  LMYADLDMDH PEVVTELKNW GKWYVNTTNI DGFRLDAVKH IKFSFFPDWL

     251                                                        300
  1  THVRNATGKE MFAVAEFWKN DLGALENYLN KTNWNHSVFD VPLHYNLYNA
  2  IHVRSATGKN MFAVAEFWKN DLGAIENYLN KTNWNHSVFD VPLHYNFYNA
  3  THVRNTTGKP MFAVAEFWKN DLGAIENYLN KTSWNHSAFD VPLHYNLYNA
  4  QAVRQATGKE MFTVAEYWQN NAGKLENYLN KTSFNQSVFD VPLHFNLQAA
  5  NHVREKTGKE MFTVAEYWQN DLGALENYLN KTNFNHSVFD VPLHYQFHAA
  6  SYVRSQTGKP LFTVGEYWSY DINKLHNYIT KTDGTMSLFD APLHNKFYTA

     301                                                        350
  1  SNSGGNYDMA KLLNGTVVQK HPMHAVTFVD NHDSQPGESL ESFVQEWFKP
  2  SKSGGNYDMR QIFNGTVVQR HPMHAVTFVD NHDSQPEEAL ESFVEEWFKP
  3  SNSGGYYDMR NILNGSVVQK HPTHAVTFVD NHDSQPGEAL ESFVQQWFKP
  4  SSQGGGYDMR RLLDGTVVSR HPEKAVTFVE NHDTQPGQSL ESTVQTWFKP
  5  STQGGGYDMR KLLNGTVVSK HPLKSVTFVD NHDTQPGQSL ESTVQTWFKP
  6  SKSGGAFDMR TLMTNTLMKD QPTLAVTFVD NHDTEPGQAL QSWVDPWFKP

     351                                                        400
  1  LAYALILTRE QGYPSVFYGD YYGIPTHS.. .VPAMKAKID PILEARQNFA
  2  LAYALTLTRE QGYPSVFYGD YYGIPTHG.. .VPAMKSKID PILEARQKYA
  3  LAYALVLTRE QGYPSVFYGD YYGIPTHG.. .VPAMKSKID PLLQARQTFA
  4  LAYAFILTRE SGYPQVFYGD MYGTKGTSPK EIPSLKDNIE PILKARKEYA
  5  LAYAFILTRE SGYPQVFYGD MYGTKGDSQR EIPALKHKIE PILKARKQYA
  6  LAYAFILTRQ EGYPCVFYGD YYGIPQYN.. .IPSLKSKID PLLIARRDYA

     401                                                        450
  1  YGTQHDYFDH HNIIGWTREG NTTHPNSGLA TIMSDGPGGE KWMYVGQNKA
  2  YGRQN..... .......... .......... .......... ..........
  3  YGTQHDYFDH HDIIGWTREG NSSHPNSGLA TIMSDGPGGN KWMYVGKNKA
  4  YGPQHDYIDH PDVIGWTREG DSSAAKSGLA ALITDGPGGS KRMYAGLKNA
  5  YGAQHDYFDH HDIVGWTREG DSSVANSGLA ALITDGPGGA KRMYVGRQNA
  6  YGTQHDYLDH SDIIGWTREG GTEKPGSGLA ALITDGPGGS KWMYVGKQHA
```

Fig. 1 (continued)

```
    451                                                     500
1   GQVWHDITGN KPGTVTINAD GWANFSVNGG SVSIWVKR.. ..........
2   .......... .......... .......... .......... ..........
3   GQVWRDITGN RTGTVTINAD GWGNFSVNGG SVSVWVKQ.. ..........
4   GETWYDITGN RSDTVKIGSD GWGEFHVNDG SVSIYVQ... ..........
5   GETWHDITGN RSEPVVINSE GWGEFHVNGG SVSIYVQR.. ..........
6   GKVFYDLTGN RSDTVTINSD GWGEFKVNGG SVSVWVPRKT TVSTIARPIT

    501            519
1   .......... .........
2   .......... .........
3   .......... .........
4   .......... .........
5   .......... .........
6   TRPWTGEFVR WTEPRLVAW
```

Fig. 1 (continued)

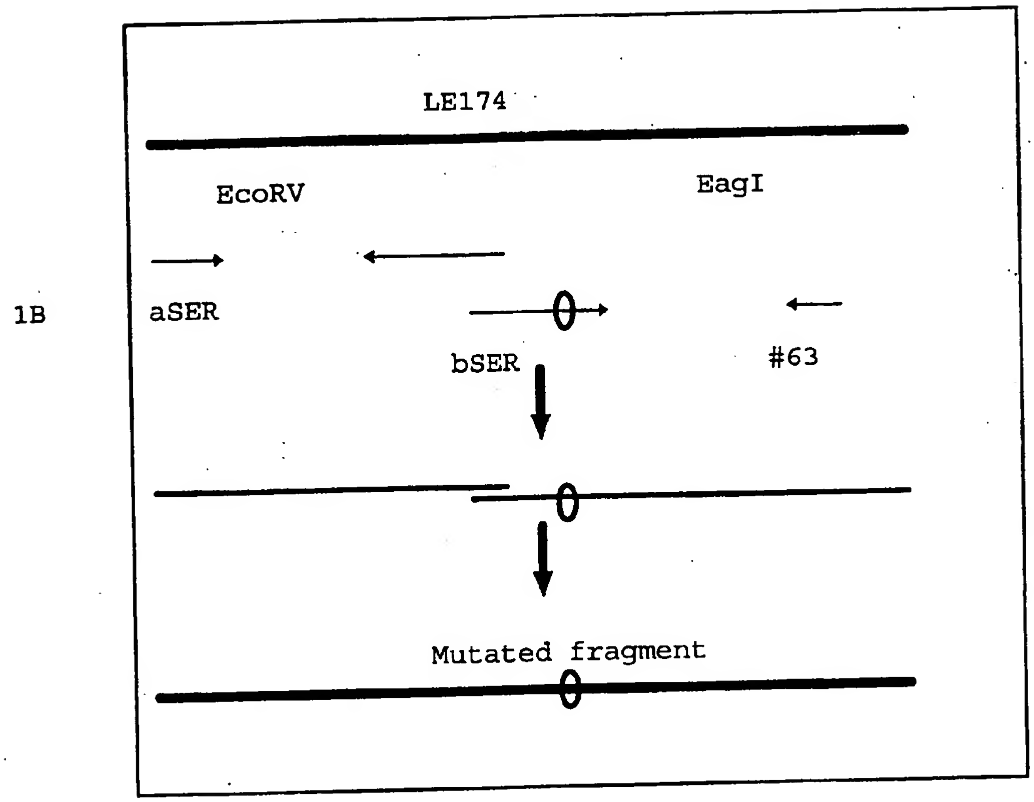

Fig. 2

SEQUENCE LISTING

(1) GENERAL INFORMATION:
(i) APPLICANT:
(A) NAME: NOVO NORDISK A/S
(B) STREET: Novo Alle
(C) CITY: DK-2880 Bagsvaerd
(E) COUNTRY: Denmark
(F) POSTAL CODE (ZIP): DK-2880
(G) TELEPHONE: +45 44 44 88 88
(H) TELEFAX: +45 44 49 32 56
(ii) TITLE OF INVENTION: α-amylase variants
(iii) NUMBER OF SEQUENCES: 32
(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS

(2) INFORMATION FOR SEQ ID NO: 1:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 485 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: protein
(iii) Organism: Bacillus sp.
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```
His His Asn Gly Thr Asn Gly Thr Met Met Gln Tyr Phe Glu Trp Tyr
1               5                   10                  15

Leu Pro Asn Asp Gly Asn His Trp Asn Arg Leu Arg Asp Asp Ala Ala
            20                  25                  30

Asn Leu Lys Ser Lys Gly Ile Thr Ala Val Trp Ile Pro Pro Ala Trp
        35                  40                  45

Lys Gly Thr Ser Gln Asn Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr
    50                  55                  60

Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly
65                  70                  75                  80

Thr Arg Asn Gln Leu Gln Ala Ala Val Thr Ser Leu Lys Asn Asn Gly
                85                  90                  95

Ile Gln Val Tyr Gly Asp Val Val Met Asn His Lys Gly Gly Ala Asp
            100                 105                 110

Gly Thr Glu Ile Val Asn Ala Val Glu Val Asn Arg Ser Asn Arg Asn
        115                 120                 125

Gln Glu Thr Ser Gly Glu Tyr Ala Ile Glu Ala Trp Thr Lys Phe Asp
    130                 135                 140

Phe Pro Gly Arg Gly Asn Asn His Ser Ser Phe Lys Trp Arg Trp Tyr
145                 150                 155                 160

His Phe Asp Gly Thr Asp Trp Asp Gln Ser Arg Gln Leu Gln Asn Lys
                165                 170                 175

Ile Tyr Lys Phe Arg Gly Thr Gly Lys Ala Trp Asp Trp Glu Val Asp
            180                 185                 190

Thr Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Val Asp Met
        195                 200                 205
```

Asp His Pro Glu Val Ile His Glu Leu Arg Asn Trp Gly Val Trp Tyr
210 215 220

Thr Asn Thr Leu Asn Leu Asp Gly Phe Arg Ile Asp Ala Val Lys His
225 230 235 240

Ile Lys Tyr Ser Phe Thr Arg Asp Trp Leu Thr His Val Arg Asn Thr
245 250 255

Thr Gly Lys Pro Met Phe Ala Val Ala Glu Phe Trp Lys Asn Asp Leu
260 265 270

Gly Ala Ile Glu Asn Tyr Leu Asn Lys Thr Ser Trp Asn His Ser Val
275 280 285

Phe Asp Val Pro Leu His Tyr Asn Leu Tyr Asn Ala Ser Asn Ser Gly
290 295 300

Gly Tyr Tyr Asp Met Arg Asn Ile Leu Asn Gly Ser Val Val Gln Lys
305 310 315 320

His Pro Thr His Ala Val Thr Phe Val Asp Asn His Asp Ser Gln Pro
325 330 335

Gly Glu Ala Leu Glu Ser Phe Val Gln Gln Trp Phe Lys Pro Leu Ala
340 345 350

Tyr Ala Leu Val Leu Thr Arg Glu Gln Gly Tyr Pro Ser Val Phe Tyr
355 360 365

Gly Asp Tyr Tyr Gly Ile Pro Thr His Gly Val Pro Ala Met Lys Ser
370 375 380

Lys Ile Asp Pro Leu Leu Gln Ala Arg Gln Thr Phe Ala Tyr Gly Thr
385 390 395 400

Gln His Asp Tyr Phe Asp His His Asp Ile Ile Gly Trp Thr Arg Glu
405 410 415

Gly Asn Ser Ser His Pro Asn Ser Gly Leu Ala Thr Ile Met Ser Asp
420 425 430

Gly Pro Gly Gly Asn Lys Trp Met Tyr Val Gly Lys Asn Lys Ala Gly
435 440 445

Gln Val Trp Arg Asp Ile Thr Gly Asn Arg Thr Gly Thr Val Thr Ile
450 455 460

Asn Ala Asp Gly Trp Gly Asn Phe Ser Val Asn Gly Gly Ser Val Ser
465 470 475 480

Val Trp Val Lys Gln
485

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 485 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) Organism: Bacillus sp.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

His His Asn Gly Thr Asn Gly Thr Met Met Gln Tyr Phe Glu Trp His
1 5 10 15

```
Leu Pro Asn Asp Gly Asn His Trp Asn Arg Leu Arg Asp Asp Ala Ser
            20                  25                  30

Asn Leu Arg Asn Arg Gly Ile Thr Ala Ile Trp Ile Pro Pro Ala Trp
        35                  40                  45

Lys Gly Thr Ser Gln Asn Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr
    50                  55                  60

Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly
65                  70                  75                  80

Thr Arg Ser Gln Leu Glu Ser Ala Ile His Ala Leu Lys Asn Asn Gly
                85                  90                  95

Val Gln Val Tyr Gly Asp Val Val Met Asn His Lys Gly Gly Ala Asp
            100                 105                 110

Ala Thr Glu Asn Val Leu Ala Val Glu Val Asn Pro Asn Asn Arg Asn
        115                 120                 125

Gln Glu Ile Ser Gly Asp Tyr Thr Ile Glu Ala Trp Thr Lys Phe Asp
    130                 135                 140

Phe Pro Gly Arg Gly Asn Thr Tyr Ser Asp Phe Lys Trp Arg Trp Tyr
145                 150                 155                 160

His Phe Asp Gly Val Asp Trp Asp Gln Ser Arg Gln Phe Gln Asn Arg
                165                 170                 175

Ile Tyr Lys Phe Arg Gly Asp Gly Lys Ala Trp Asp Trp Glu Val Asp
            180                 185                 190

Ser Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Val Asp Met
        195                 200                 205

Asp His Pro Glu Val Val Asn Glu Leu Arg Arg Trp Gly Glu Trp Tyr
    210                 215                 220

Thr Asn Thr Leu Asn Leu Asp Gly Phe Arg Ile Asp Ala Val Lys His
225                 230                 235                 240

Ile Lys Tyr Ser Phe Thr Arg Asp Trp Leu Thr His Val Arg Asn Ala
                245                 250                 255

Thr Gly Lys Glu Met Phe Ala Val Ala Glu Phe Trp Lys Asn Asp Leu
            260                 265                 270

Gly Ala Leu Glu Asn Tyr Leu Asn Lys Thr Asn Trp Asn His Ser Val
        275                 280                 285

Phe Asp Val Pro Leu His Tyr Asn Leu Tyr Asn Ala Ser Asn Ser Gly
    290                 295                 300

Gly Asn Tyr Asp Met Ala Lys Leu Leu Asn Gly Thr Val Val Gln Lys
305                 310                 315                 320

His Pro Met His Ala Val Thr Phe Val Asp Asn His Asp Ser Gln Pro
                325                 330                 335

Gly Glu Ser Leu Glu Ser Phe Val Gln Glu Trp Phe Lys Pro Leu Ala
            340                 345                 350

Tyr Ala Leu Ile Leu Thr Arg Glu Gln Gly Tyr Pro Ser Val Phe Tyr
        355                 360                 365

Gly Asp Tyr Tyr Gly Ile Pro Thr His Ser Val Pro Ala Met Lys Ala
```

370 375 380

Lys Ile Asp Pro Ile Leu Glu Ala Arg Gln Asn Phe Ala Tyr Gly Thr
385 390 395 400

Gln His Asp Tyr Phe Asp His His Asn Ile Ile Gly Trp Thr Arg Glu
405 410 415

Gly Asn Thr Thr His Pro Asn Ser Gly Leu Ala Thr Ile Met Ser Asp
420 425 430

Gly Pro Gly Gly Glu Lys Trp Met Tyr Val Gly Gln Asn Lys Ala Gly
435 440 445

Gln Val Trp His Asp Ile Thr Gly Asn Lys Pro Gly Thr Val Thr Ile
450 455 460

Asn Ala Asp Gly Trp Ala Asn Phe Ser Val Asn Gly Gly Ser Val Ser
465 470 475 480

Ile Trp Val Lys Arg
485

(2) INFORMATION FOR SEQ ID NO: 3:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 514 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: protein
(iii) Organism: Bacillus stearothermophilus.
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Ala Ala Pro Phe Asn Gly Thr Met Met Gln Tyr Phe Glu Trp Tyr Leu
1 5 10 15

Pro Asp Asp Gly Thr Leu Trp Thr Lys Val Ala Asn Glu Ala Asn Asn
20 25 30

Leu Ser Ser Leu Gly Ile Thr Ala Leu Trp Leu Pro Pro Ala Tyr Lys
35 40 45

Gly Thr Ser Arg Ser Asp Val Gly Tyr Gly Val Tyr Asp Leu Tyr Asp
50 55 60

Leu Gly Glu Phe Asn Gln Lys Gly Ala Val Arg Thr Lys Tyr Gly Thr
65 70 75 80

Lys Ala Gln Tyr Leu Gln Ala Ile Gln Ala Ala His Ala Ala Gly Met
85 90 95

Gln Val Tyr Ala Asp Val Val Phe Asp His Lys Gly Gly Ala Asp Gly
100 105 110

Thr Glu Trp Val Asp Ala Val Glu Val Asn Pro Ser Asp Arg Asn Gln
115 120 125

Glu Ile Ser Gly Thr Tyr Gln Ile Gln Ala Trp Thr Lys Phe Asp Phe
130 135 140

Pro Gly Arg Gly Asn Thr Tyr Ser Ser Phe Lys Trp Arg Trp Tyr His
145 150 155 160

Phe Asp Gly Val Asp Trp Asp Glu Ser Arg Lys Leu Ser Arg Ile Tyr
165 170 175

Lys Phe Arg Gly Ile Gly Lys Ala Trp Asp Trp Glu Val Asp Thr Glu

180 185 190

Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Leu Asp Met Asp His
195 200 205

Pro Glu Val Val Thr Glu Leu Lys Ser Trp Gly Lys Trp Tyr Val Asn
210 215 220

Thr Thr Asn Ile Asp Gly Phe Arg Leu Asp Ala Val Lys His Ile Lys
225 230 235 240

Phe Ser Phe Phe Pro Asp Trp Leu Ser Asp Val Arg Ser Gln Thr Gly
245 250 255

Lys Pro Leu Phe Thr Val Gly Glu Tyr Trp Ser Tyr Asp Ile Asn Lys
260 265 270

Leu His Asn Tyr Ile Met Lys Thr Asn Gly Thr Met Ser Leu Phe Asp
275 280 285

Ala Pro Leu His Asn Lys Phe Tyr Thr Ala Ser Lys Ser Gly Gly Thr
290 295 300

Phe Asp Met Arg Thr Leu Met Thr Asn Thr Leu Met Lys Asp Gln Pro
305 310 315 320

Thr Leu Ala Val Thr Phe Val Asp Asn His Asp Thr Glu Pro Gly Gln
325 330 335

Ala Leu Gln Ser Trp Val Asp Pro Trp Phe Lys Pro Leu Ala Tyr Ala
340 345 350

Phe Ile Leu Thr Arg Gln Glu Gly Tyr Pro Cys Val Phe Tyr Gly Asp
355 360 365

Tyr Tyr Gly Ile Pro Gln Tyr Asn Ile Pro Ser Leu Lys Ser Lys Ile
370 375 380

Asp Pro Leu Leu Ile Ala Arg Arg Asp Tyr Ala Tyr Gly Thr Gln His
385 390 395 400

Asp Tyr Leu Asp His Ser Asp Ile Ile Gly Trp Thr Arg Glu Gly Val
405 410 415

Thr Glu Lys Pro Gly Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly Pro
420 425 430

Gly Gly Ser Lys Trp Met Tyr Val Gly Lys Gln His Ala Gly Lys Val
435 440 445

Phe Tyr Asp Leu Thr Gly Asn Arg Ser Asp Thr Val Thr Ile Asn Ser
450 455 460

Asp Gly Trp Gly Glu Phe Lys Val Asn Gly Gly Ser Val Ser Val Trp
465 470 475 480

Val Pro Arg Lys Thr Thr Val Ser Thr Ile Ala Trp Ser Ile Thr Thr
485 490 495

Arg Pro Trp Thr Asp Glu Phe Val Arg Trp Thr Glu Pro Arg Leu Val
500 505 510

Ala Trp

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 483 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: protein
(iii)Organism: Bacillus licheniformis
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Ala Asn Leu Asn Gly Thr Leu Met Gln Tyr Phe Glu Trp Tyr Met Pro
1 5 10 15

Asn Asp Gly Gln His Trp Arg Arg Leu Gln Asn Asp Ser Ala Tyr Leu
20 25 30

Ala Glu His Gly Ile Thr Ala Val Trp Ile Pro Pro Ala Tyr Lys Gly
35 40 45

Thr Ser Gln Ala Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr Asp Leu
50 55 60

Gly Glu Phe His Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr Lys
65 70 75 80

Gly Glu Leu Gln Ser Ala Ile Lys Ser Leu His Ser Arg Asp Ile Asn
85 90 95

Val Tyr Gly Asp Val Val Ile Asn His Lys Gly Gly Ala Asp Ala Thr
100 105 110

Glu Asp Val Thr Ala Val Glu Val Asp Pro Ala Asp Arg Asn Arg Val
115 120 125

Ile Ser Gly Glu His Leu Ile Lys Ala Trp Thr His Phe His Phe Pro
130 135 140

Gly Arg Gly Ser Thr Tyr Ser Asp Phe Lys Trp His Trp Tyr His Phe
145 150 155 160

Asp Gly Thr Asp Trp Asp Glu Ser Arg Lys Leu Asn Arg Ile Tyr Lys
165 170 175

Phe Gln Gly Lys Ala Trp Asp Trp Glu Val Ser Asn Glu Asn Gly Asn
180 185 190

Tyr Asp Tyr Leu Met Tyr Ala Asp Ile Asp Tyr Asp His Pro Asp Val
195 200 205

Ala Ala Glu Ile Lys Arg Trp Gly Thr Trp Tyr Ala Asn Glu Leu Gln
210 215 220

Leu Asp Gly Phe Arg Leu Asp Ala Val Lys His Ile Lys Phe Ser Phe
225 230 235 240

Leu Arg Asp Trp Val Asn His Val Arg Glu Lys Thr Gly Lys Glu Met
245 250 255

Phe Thr Val Ala Glu Tyr Trp Gln Asn Asp Leu Gly Ala Leu Glu Asn
260 265 270

Tyr Leu Asn Lys Thr Asn Phe Asn His Ser Val Phe Asp Val Pro Leu
275 280 285

His Tyr Gln Phe His Ala Ala Ser Thr Gln Gly Gly Gly Tyr Asp Met
290 295 300

Arg Lys Leu Leu Asn Gly Thr Val Val Ser Lys His Pro Leu Lys Ser
305 310 315 320

```
Val Thr Phe Val Asp Asn His Asp Thr Gln Pro Gly Gln Ser Leu Glu
                325                 330                 335

Ser Thr Val Gln Thr Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu
            340                 345                 350

Thr Arg Glu Ser Gly Tyr Pro Gln Val Phe Tyr Gly Asp Met Tyr Gly
        355                 360                 365

Thr Lys Gly Asp Ser Gln Arg Glu Ile Pro Ala Leu Lys His Lys Ile
    370                 375                 380

Glu Pro Ile Leu Lys Ala Arg Lys Gln Tyr Ala Tyr Gly Ala Gln His
385                 390                 395                 400

Asp Tyr Phe Asp His His Asp Ile Val Gly Trp Thr Arg Glu Gly Asp
                405                 410                 415

Ser Ser Val Ala Asn Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly Pro
            420                 425                 430

Gly Gly Ala Lys Arg Met Tyr Val Gly Arg Gln Asn Ala Gly Glu Thr
        435                 440                 445

Trp His Asp Ile Thr Gly Asn Arg Ser Glu Pro Val Val Ile Asn Ser
    450                 455                 460

Glu Gly Trp Gly Glu Phe His Val Asn Gly Gly Ser Val Ser Ile Tyr
465                 470                 475                 480

Val Gln Arg
```

(2) INFORMATION FOR SEQ ID NO: 5:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 480 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: protein
(iii) Organism: Bacillus amyloliquefaciens
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

```
Val Asn Gly Thr Leu Met Gln Tyr Phe Glu Trp Tyr Thr Pro Asn Asp
1               5                   10                  15

Gly Gln His Trp Lys Arg Leu Gln Asn Asp Ala Glu His Leu Ser Asp
            20                  25                  30

Ile Gly Ile Thr Ala Val Trp Ile Pro Pro Ala Tyr Lys Gly Leu Ser
        35                  40                  45

Gln Ser Asp Asn Gly Tyr Gly Pro Tyr Asp Leu Tyr Asp Leu Gly Glu
    50                  55                  60

Phe Gln Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr Lys Ser Glu
65                  70                  75                  80

Leu Gln Asp Ala Ile Gly Ser Leu His Ser Arg Asn Val Gln Val Tyr
                85                  90                  95

Gly Asp Val Val Leu Asn His Lys Ala Gly Ala Asp Ala Thr Glu Asp
            100                 105                 110

Val Thr Ala Val Glu Val Asn Pro Ala Asn Arg Asn Gln Glu Thr Ser
        115                 120                 125
```

```
Glu Glu Tyr Gln Ile Lys Ala Trp Thr Asp Phe Arg Phe Pro Gly Arg
    130                 135                 140

Gly Asn Thr Tyr Ser Asp Phe Lys Trp His Trp Tyr His Phe Asp Gly
145                 150                 155                 160

Ala Asp Trp Asp Glu Ser Arg Lys Ile Ser Arg Ile Phe Lys Phe Arg
                165                 170                 175

Gly Glu Gly Lys Ala Trp Asp Trp Glu Val Ser Ser Glu Asn Gly Asn
            180                 185                 190

Tyr Asp Tyr Leu Met Tyr Ala Asp Val Asp Tyr Asp His Pro Asp Val
        195                 200                 205

Val Ala Glu Thr Lys Lys Trp Gly Ile Trp Tyr Ala Asn Glu Leu Ser
    210                 215                 220

Leu Asp Gly Phe Arg Ile Asp Ala Ala Lys His Ile Lys Phe Ser Phe
225                 230                 235                 240

Leu Arg Asp Trp Val Gln Ala Val Arg Gln Ala Thr Gly Lys Glu Met
                245                 250                 255

Phe Thr Val Ala Glu Tyr Trp Gln Asn Asn Ala Gly Lys Leu Glu Asn
            260                 265                 270

Tyr Leu Asn Lys Thr Ser Phe Asn Gln Ser Val Phe Asp Val Pro Leu
        275                 280                 285

His Phe Asn Leu Gln Ala Ala Ser Ser Gln Gly Gly Gly Tyr Asp Met
    290                 295                 300

Arg Arg Leu Leu Asp Gly Thr Val Val Ser Arg His Pro Glu Lys Ala
305                 310                 315                 320

Val Thr Phe Val Glu Asn His Asp Thr Gln Pro Gly Gln Ser Leu Glu
                325                 330                 335

Ser Thr Val Gln Thr Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu
            340                 345                 350

Thr Arg Glu Ser Gly Tyr Pro Gln Val Phe Tyr Gly Asp Met Tyr Gly
        355                 360                 365

Thr Lys Gly Thr Ser Pro Lys Glu Ile Pro Ser Leu Lys Asp Asn Ile
    370                 375                 380

Glu Pro Ile Leu Lys Ala Arg Lys Glu Tyr Ala Tyr Gly Pro Gln His
385                 390                 395                 400

Asp Tyr Ile Asp His Pro Asp Val Ile Gly Trp Thr Arg Glu Gly Asp
                405                 410                 415

Ser Ser Ala Ala Lys Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly Pro
            420                 425                 430

Gly Gly Ser Lys Arg Met Tyr Ala Gly Leu Lys Asn Ala Gly Glu Thr
        435                 440                 445

Trp Tyr Asp Ile Thr Gly Asn Arg Ser Asp Thr Val Lys Ile Gly Ser
    450                 455                 460

Asp Gly Trp Gly Glu Phe His Val Asn Asp Gly Ser Val Ser Ile Tyr
465                 470                 475                 480
```

(2) INFORMATION FOR SEQ ID NO: 6:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 485 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: protein
(iii)Organism: Bacillus sp.
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

```
His His Asn Gly Thr Asn Gly Thr Met Met Gln Tyr Phe Glu Trp Tyr
1               5                   10                  15

Leu Pro Asn Asp Gly Asn His Trp Asn Arg Leu Asn Ser Asp Ala Ser
            20                  25                  30

Asn Leu Lys Ser Lys Gly Ile Thr Ala Val Trp Ile Pro Pro Ala Trp
        35                  40                  45

Lys Gly Ala Ser Gln Asn Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr
    50                  55                  60

Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly
65                  70                  75                  80

Thr Arg Ser Gln Leu Gln Ala Ala Val Thr Ser Leu Lys Asn Asn Gly
                85                  90                  95

Ile Gln Val Tyr Gly Asp Val Val Met Asn His Lys Gly Gly Ala Asp
            100                 105                 110

Ala Thr Glu Met Val Arg Ala Val Glu Val Asn Pro Asn Asn Arg Asn
        115                 120                 125

Gln Glu Val Thr Gly Glu Tyr Thr Ile Glu Ala Trp Thr Arg Phe Asp
    130                 135                 140

Phe Pro Gly Arg Gly Asn Thr His Ser Ser Phe Lys Trp Arg Trp Tyr
145                 150                 155                 160

His Phe Asp Gly Val Asp Trp Asp Gln Ser Arg Arg Leu Asn Asn Arg
                165                 170                 175

Ile Tyr Lys Phe Arg Gly His Gly Lys Ala Trp Asp Trp Glu Val Asp
            180                 185                 190

Thr Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Ile Asp Met
        195                 200                 205

Asp His Pro Glu Val Val Asn Glu Leu Arg Asn Trp Gly Val Trp Tyr
    210                 215                 220

Thr Asn Thr Leu Gly Leu Asp Gly Phe Arg Ile Asp Ala Val Lys His
225                 230                 235                 240

Ile Lys Tyr Ser Phe Thr Arg Asp Trp Ile Asn His Val Arg Ser Ala
                245                 250                 255

Thr Gly Lys Asn Met Phe Ala Val Ala Glu Phe Trp Lys Asn Asp Leu
            260                 265                 270

Gly Ala Ile Glu Asn Tyr Leu Gln Lys Thr Asn Trp Asn His Ser Val
        275                 280                 285
```

```
Phe Asp Val Pro Leu His Tyr Asn Leu Tyr Asn Ala Ser Lys Ser Gly
    290                 295                 300

Gly Asn Tyr Asp Met Arg Asn Ile Phe Asn Gly Thr Val Val Gln Arg
305                 310                 315                 320

His Pro Ser His Ala Val Thr Phe Val Asp Asn His Asp Ser Gln Pro
                325                 330                 335

Glu Glu Ala Leu Glu Ser Phe Val Glu Glu Trp Phe Lys Pro Leu Ala
            340                 345                 350

Tyr Ala Leu Thr Leu Thr Arg Glu Gln Gly Tyr Pro Ser Val Phe Tyr
        355                 360                 365

Gly Asp Tyr Tyr Gly Ile Pro Thr His Gly Val Pro Ala Met Arg Ser
    370                 375                 380

Lys Ile Asp Pro Ile Leu Glu Ala Arg Gln Lys Tyr Ala Tyr Gly Lys
385                 390                 395                 400

Gln Asn Asp Tyr Leu Asp His His Asn Ile Ile Gly Trp Thr Arg Glu
                405                 410                 415

Gly Asn Thr Ala His Pro Asn Ser Gly Leu Ala Thr Ile Met Ser Asp
            420                 425                 430

Gly Ala Gly Gly Ser Lys Trp Met Phe Val Gly Arg Asn Lys Ala Gly
        435                 440                 445

Gln Val Trp Ser Asp Ile Thr Gly Asn Arg Thr Gly Thr Val Thr Ile
    450                 455                 460

Asn Ala Asp Gly Trp Gly Asn Phe Ser Val Asn Gly Gly Ser Val Ser
465                 470                 475                 480

Ile Trp Val Asn Lys
                485
```

(2) INFORMATION FOR SEQ ID NO: 7:
  (i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 485 amino acids
    (B) TYPE: amino acid
    (C) STRANDEDNESS: single
    (D) TOPOLOGY: linear
  (ii) MOLECULE TYPE: protein
  (iii) Organism: Bacillus sp.
  (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

```
His His Asn Gly Thr Asn Gly Thr Met Met Gln Tyr Phe Glu Trp Tyr
1               5                   10                  15

Leu Pro Asn Asp Gly Asn His Trp Asn Arg Leu Arg Asp Asp Ala Ala
            20                  25                  30

Asn Leu Lys Ser Lys Gly Ile Thr Ala Val Trp Ile Pro Pro Ala Trp
        35                  40                  45

Lys Gly Thr Ser Gln Asn Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr
    50                  55                  60

Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly
65                  70                  75                  80

Thr Arg Asn Gln Leu Gln Ala Ala Val Thr Ser Leu Lys Asn Asn Gly
                85                  90                  95
```

```
Ile Gln Val Tyr Gly Asp Val Val Met Asn His Lys Gly Gly Ala Asp
            100                 105                 110

Gly Thr Glu Ile Val Asn Ala Val Glu Val Asn Arg Ser Asn Arg Asn
        115                 120                 125

Gln Glu Thr Ser Gly Glu Tyr Ala Ile Glu Ala Trp Thr Lys Phe Asp
    130                 135                 140

Phe Pro Gly Arg Gly Asn Asn His Ser Ser Phe Lys Trp Arg Trp Tyr
145                 150                 155                 160

His Phe Asp Gly Thr Asp Trp Asp Gln Ser Arg Gln Leu Gln Asn Lys
                165                 170                 175

Ile Tyr Lys Phe Arg Gly Thr Gly Lys Ala Trp Asp Trp Glu Val Asp
            180                 185                 190

Thr Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Val Asp Met
        195                 200                 205

Asp His Pro Glu Val Ile His Glu Leu Arg Asn Trp Gly Val Trp Tyr
    210                 215                 220

Thr Asn Thr Leu Asn Leu Asp Gly Phe Arg Ile Asp Ala Val Lys His
225                 230                 235                 240

Ile Lys Tyr Ser Phe Thr Arg Asp Trp Leu Thr His Val Arg Asn Thr
                245                 250                 255

Thr Gly Lys Pro Met Phe Ala Val Ala Glu Phe Trp Lys Asn Asp Leu
            260                 265                 270

Gly Ala Ile Glu Asn Tyr Leu Asn Lys Thr Ser Trp Asn His Ser Val
        275                 280                 285

Phe Asp Val Pro Leu His Tyr Asn Leu Tyr Asn Ala Ser Asn Ser Gly
    290                 295                 300

Gly Tyr Tyr Asp Met Arg Asn Ile Leu Asn Gly Ser Val Val Gln Lys
305                 310                 315                 320

His Pro Thr His Ala Val Thr Phe Val Asp Asn His Asp Ser Gln Pro
                325                 330                 335

Gly Glu Ala Leu Glu Ser Phe Val Gln Gln Trp Phe Lys Pro Leu Ala
            340                 345                 350

Tyr Ala Leu Val Leu Thr Arg Glu Gln Gly Tyr Pro Ser Val Phe Tyr
        355                 360                 365

Gly Asp Tyr Tyr Gly Ile Pro Thr His Gly Val Pro Ala Met Lys Ser
    370                 375                 380

Lys Ile Asp Pro Leu Leu Gln Ala Arg Gln Thr Phe Ala Tyr Gly Thr
385                 390                 395                 400

Gln His Asp Tyr Phe Asp His His Asp Ile Ile Gly Trp Thr Arg Glu
                405                 410                 415

Gly Asn Ser Ser His Pro Asn Ser Gly Leu Ala Thr Ile Met Ser Asp
            420                 425                 430

Gly Pro Gly Gly Asn Lys Trp Met Tyr Val Gly Lys Asn Lys Ala Gly
        435                 440                 445

Gln Val Trp Arg Asp Ile Thr Gly Asn Arg Thr Gly Thr Val Thr Ile
```

450 455 460

Asn Ala Asp Gly Trp Gly Asn Phe Ser Val Asn Gly Gly Ser Val Ser
465 470 475 480

Val Trp Val Lys Gln
485

(2) INFORMATION FOR SEQ ID NO: 8:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 485 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: protein
(iii) Organism: Bacillus sp.
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

His His Asn Gly Thr Asn Gly Thr Met Met Gln Tyr Phe Glu Trp His
1 5 10 15

Leu Pro Asn Asp Gly Asn His Trp Asn Arg Leu Arg Asp Asp Ala Ser
20 25 30

Asn Leu Arg Asn Arg Gly Ile Thr Ala Ile Trp Ile Pro Pro Ala Trp
35 40 45

Lys Gly Thr Ser Gln Asn Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr
50 55 60

Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly
65 70 75 80

Thr Arg Ser Gln Leu Glu Ser Ala Ile His Ala Leu Lys Asn Asn Gly
85 90 95

Val Gln Val Tyr Gly Asp Val Val Met Asn His Lys Gly Gly Ala Asp
100 105 110

Ala Thr Glu Asn Val Leu Ala Val Glu Val Asn Pro Asn Asn Arg Asn
115 120 125

Gln Glu Ile Ser Gly Asp Tyr Thr Ile Glu Ala Trp Thr Lys Phe Asp
130 135 140

Phe Pro Gly Arg Gly Asn Thr Tyr Ser Asp Phe Lys Trp Arg Trp Tyr
145 150 155 160

His Phe Asp Gly Val Asp Trp Asp Gln Ser Arg Gln Phe Gln Asn Arg
165 170 175

Ile Tyr Lys Phe Arg Gly Asp Gly Lys Ala Trp Asp Trp Glu Val Asp
180 185 190

Ser Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Val Asp Met
195 200 205

Asp His Pro Glu Val Val Asn Glu Leu Arg Arg Trp Gly Glu Trp Tyr
210 215 220

Thr Asn Thr Leu Asn Leu Asp Gly Phe Arg Ile Asp Ala Val Lys His
225 230 235 240

Ile Lys Tyr Ser Phe Thr Arg Asp Trp Leu Thr His Val Arg Asn Ala
245 250 255

Thr Gly Lys Glu Met Phe Ala Val Ala Glu Phe Trp Lys Asn Asp Leu

260 265 270

Gly Ala Leu Glu Asn Tyr Leu Asn Lys Thr Asn Trp Asn His Ser Val
275 280 285

Phe Asp Val Pro Leu His Tyr Asn Leu Tyr Asn Ala Ser Asn Ser Gly
290 295 300

Gly Asn Tyr Asp Met Ala Lys Leu Leu Asn Gly Thr Val Val Gln Lys
305 310 315 320

His Pro Met His Ala Val Thr Phe Val Asp Asn His Asp Ser Gln Pro
325 330 335

Gly Glu Ser Leu Glu Ser Phe Val Gln Glu Trp Phe Lys Pro Leu Ala
340 345 350

Tyr Ala Leu Ile Leu Thr Arg Glu Gln Gly Tyr Pro Ser Val Phe Tyr
355 360 365

Gly Asp Tyr Tyr Gly Ile Pro Thr His Ser Val Pro Ala Met Lys Ala
370 375 380

Lys Ile Asp Pro Ile Leu Glu Ala Arg Gln Asn Phe Ala Tyr Gly Thr
385 390 395 400

Gln His Asp Tyr Phe Asp His His Asn Ile Ile Gly Trp Thr Arg Glu
405 410 415

Gly Asn Thr Thr His Pro Asn Ser Gly Leu Ala Thr Ile Met Ser Asp
420 425 430

Gly Pro Gly Gly Glu Lys Trp Met Tyr Val Gly Gln Asn Lys Ala Gly
435 440 445

Gln Val Trp His Asp Ile Thr Gly Asn Lys Pro Gly Thr Val Thr Ile
450 455 460

Asn Ala Asp Gly Trp Ala Asn Phe Ser Val Asn Gly Gly Ser Val Ser
465 470 475 480

Ile Trp Val Lys Arg
485

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1455 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii)Organism: Bacillus sp.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

CATCATAATG GAACAAATGG TACTATGATG CAATATTTCG AATGGTATTT GCCAAATGAC 60

GGGAATCATT GGAACAGGTT GAGGGATGAC GCAGCTAACT TAAAGAGTAA AGGGATAACA 120

GCTGTATGGA TCCCACCTGC ATGGAAGGGG ACTTCCCAGA ATGATGTAGG TTATGGAGCC 180

TATGATTTAT ATGATCTTGG AGAGTTTAAC CAGAAGGGGA CGGTTCGTAC AAAATATGGA 240

ACACGCAACC AGCTACAGGC TGCGGTGACC TCTTTAAAAA ATAACGGCAT TCAGGTATAT 300

GGTGATGTCG TCATGAATCA TAAAGGTGGA GCAGATGGTA CGGAAATTGT AAATGCGGTA 360

GAAGTGAATC GGAGCAACCG AAACCAGGAA ACCTCAGGAG AGTATGCAAT AGAAGCGTGG 420

ACAAAGTTTG ATTTTCCTGG AAGAGGAAAT AACCATTCCA GCTTTAAGTG GCGCTGGTAT 480

CATTTTGATG GGACAGATTG GGATCAGTCA CGCCAGCTTC AAAACAAAAT ATATAAATTC 540

AGGGGAACAG GCAAGGCCTG GGACTGGGAA GTCGATACAG AGAATGGCAA CTATGACTAT 600

CTTATGTATG CAGACGTGGA TATGGATCAC CCAGAAGTAA TACATGAACT TAGAAACTGG 660

GGAGTGTGGT ATACGAATAC ACTGAACCTT GATGGATTTA GAATAGATGC AGTGAAACAT 720

ATAAAATATA GCTTTACGAG AGATTGGCTT ACACATGTGC GTAACACCAC AGGTAAACCA 780

ATGTTTGCAG TGGCTGAGTT TTGGAAAAAT GACCTTGGTG CAATTGAAAA CTATTTGAAT 840

AAAACAAGTT GGAATCACTC GGTGTTTGAT GTTCCTCTCC ACTATAATTT GTACAATGCA 900

TCTAATAGCG GTGGTTATTA TGATATGAGA AATATTTTAA ATGGTTCTGT GGTGCAAAAA 960

CATCCAACAC ATGCCGTTAC TTTTGTTGAT AACCATGATT CTCAGCCCGG GGAAGCATTG 1020

GAATCCTTTG TTCAACAATG GTTTAAACCA CTTGCATATG CATTGGTTCT GACAAGGGAA 1080

CAAGGTTATC CTTCCGTATT TTATGGGGAT TACTACGGTA TCCCAACCCA TGGTGTTCCG 1140

GCTATGAAAT CTAAAATAGA CCCTCTTCTG CAGGCACGTC AAACTTTTGC CTATGGTACG 1200

CAGCATGATT ACTTTGATCA TCATGATATT ATCGGTTGGA CAAGAGAGGG AAATAGCTCC 1260

CATCCAAATT CAGGCCTTGC CACCATTATG TCAGATGGTC CAGGTGGTAA CAAATGGATG 1320

TATGTGGGGA AAAATAAAGC GGGACAAGTT TGGAGAGATA TTACCGGAAA TAGGACAGGC 1380

ACCGTCACAA TTAATGCAGA CGGATGGGGT AATTTCTCTG TTAATGGAGG GTCCGTTTCG 1440

GTTTGGGTGA AGCAA 1455

(2) INFORMATION FOR SEQ ID NO: 10:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1455 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: DNA (genomic)
(iii)Organism: Bacillus sp.
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

CATCATAATG GGACAAATGG GACGATGATG CAATACTTTG AATGGCACTT GCCTAATGAT 60

GGGAATCACT GGAATAGATT AAGAGATGAT GCTAGTAATC TAAGAAATAG AGGTATAACC 120

GCTATTTGGA TTCCGCCTGC CTGGAAAGGG ACTTCGCAAA ATGATGTGGG GTATGGAGCC 180

TATGATCTTT ATGATTTAGG GGAATTTAAT CAAAAGGGGA CGGTTCGTAC TAAGTATGGG 240

ACACGTAGTC AATTGGAGTC TGCCATCCAT GCTTTAAAGA ATAATGGCGT TCAAGTTTAT 300

GGGGATGTAG TGATGAACCA TAAAGGAGGA GCTGATGCTA CAGAAAACGT TCTTGCTGTC 360

GAGGTGAATC CAAATAACCG GAATCAAGAA ATATCTGGGG ACTACACAAT TGAGGCTTGG 420

ACTAAGTTTG ATTTTCCAGG GAGGGGTAAT ACATACTCAG ACTTTAAATG GCGTTGGTAT 480

CATTTCGATG GTGTAGATTG GGATCAATCA CGACAATTCC AAAATCGTAT CTACAAATTC 540

CGAGGTGATG GTAAGGCATG GGATTGGGAA GTAGATTCGG AAAATGGAAA TTATGATTAT 600

```
TTAATGTATG CAGATGTAGA TATGGATCAT CCGGAGGTAG TAAATGAGCT TAGAAGATGG     660
GGAGAATGGT ATACAAATAC ATTAAATCTT GATGGATTTA GGATCGATGC GGTGAAGCAT     720
ATTAAATATA GCTTTACACG TGATTGGTTG ACCCATGTAA GAAACGCAAC GGGAAAAGAA     780
ATGTTTGCTG TTGCTGAATT TTGGAAAAAT GATTTAGGTG CCTTGGAGAA CTATTTAAAT     840
AAAACAAACT GGAATCATTC TGTCTTTGAT GTCCCCCTTC ATTATAATCT TTATAACGCG     900
TCAAATAGTG GAGGCAACTA TGACATGGCA AAACTTCTTA ATGGAACGGT TGTTCAAAAG     960
CATCCAATGC ATGCCGTAAC TTTTGTGGAT AATCACGATT CTCAACCTGG GGAATCATTA    1020
GAATCATTTG TACAAGAATG GTTTAAGCCA CTTGCTTATG CGCTTATTTT AACAAGAGAA    1080
CAAGGCTATC CCTCTGTCTT CTATGGTGAC TACTATGGAA TTCCAACACA TAGTGTCCCA    1140
GCAATGAAAG CCAAGATTGA TCCAATCTTA GAGGCGCGTC AAAATTTTGC ATATGGAACA    1200
CAACATGATT ATTTTGACCA TCATAATATA ATCGGATGGA CACGTGAAGG AAATACCACG    1260
CATCCCAATT CAGGACTTGC GACTATCATG TCGGATGGGC CAGGGGGAGA GAAATGGATG    1320
TACGTAGGGC AAAATAAAGC AGGTCAAGTT TGGCATGACA TAACTGGAAA TAAACCAGGA    1380
ACAGTTACGA TCAATGCAGA TGGATGGGCT AATTTTTCAG TAAATGGAGG ATCTGTTTCC    1440
ATTTGGGTGA AACGA                                                     1455
```

(2) INFORMATION FOR SEQ ID NO: 11:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1548 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: DNA (genomic)
(iii)Organism: Bacillus stearothermophilus
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

```
GCCGCACCGT TTAACGGCAC CATGATGCAG TATTTTGAAT GGTACTTGCC GGATGATGGC      60
ACGTTATGGA CCAAAGTGGC CAATGAAGCC AACAACTTAT CCAGCCTTGG CATCACCGCT     120
CTTTGGCTGC CGCCCGCTTA CAAAGGAACA AGCCGCAGCG ACGTAGGGTA CGGAGTATAC     180
GACTTGTATG ACCTCGGCGA ATTCAATCAA AAAGGGACCG TCCGCACAAA ATACGGAACA     240
AAAGCTCAAT ATCTTCAAGC CATTCAAGCC GCCCACGCCG CTGGAATGCA AGTGTACGCC     300
GATGTCGTGT TCGACCATAA AGGCGGCGCT GACGGCACGG AATGGGTGGA CGCCGTCGAA     360
GTCAATCCGT CCGACCGCAA CCAAGAAATC TCGGGCACCT ATCAAATCCA AGCATGGACG     420
AAATTTGATT TTCCCGGGCG GGGCAACACC TACTCCAGCT TTAAGTGGCG CTGGTACCAT     480
TTTGACGGCG TTGATTGGGA CGAAAGCCGA AAATTGAGCC GCATTTACAA ATTCCGCGGC     540
ATCGGCAAAG CGTGGGATTG GGAAGTAGAC ACGGAAAACG GAAACTATGA CTACTTAATG     600
TATGCCGACC TTGATATGGA TCATCCCGAA GTCGTGACCG AGCTGAAAAA CTGGGGGAAA     660
TGGTATGTCA ACACAACGAA CATTGATGGG TTCCGGCTTG ATGCCGTCAA GCATATTAAG     720
TTCAGTTTTT TTCCTGATTG GTTGTCGTAT GTGCGTTCTC AGACTGGCAA GCCGCTATTT     780
```

ACCGTCGGGG AATATTGGAG CTATGACATC AACAAGTTGC ACAATTACAT TACGAAAACA 840

GACGGAACGA TGTCTTTGTT TGATGCCCCG TTACACAACA AATTTTATAC CGCTTCCAAA 900

TCAGGGGGCG CATTTGATAT GCGCACGTTA ATGACCAATA CTCTCATGAA AGATCAACCG 960

ACATTGGCCG TCACCTTCGT TGATAATCAT GACACCGAAC CCGGCCAAGC GCTGCAGTCA 1020

TGGGTCGACC CATGGTTCAA ACCGTTGGCT TACGCCTTTA TTCTAACTCG GCAGGAAGGA 1080

TACCCGTGCG TCTTTTATGG TGACTATTAT GGCATTCCAC AATATAACAT TCCTTCGCTG 1140

AAAAGCAAAA TCGATCCGCT CCTCATCGCG CGCAGGGATT ATGCTTACGG AACGCAACAT 1200

GATTATCTTG ATCACTCCGA CATCATCGGG TGGACAAGGG AAGGGGGCAC TGAAAAACCA 1260

GGATCCGGAC TGGCCGCACT GATCACCGAT GGGCCGGGAG GAAGCAAATG GATGTACGTT 1320

GGCAAACAAC ACGCTGGAAA AGTGTTCTAT GACCTTACCG GCAACCGGAG TGACACCGTC 1380

ACCATCAACA GTGATGGATG GGGGGAATTC AAAGTCAATG GCGGTTCGGT TTCGGTTTGG 1440

GTTCCTAGAA AAACGACCGT TTCTACCATC GCTCGGCCGA TCACAACCCG ACCGTGGACT 1500

GGTGAATTCG TCCGTTGGAC CGAACCACGG TTGGTGGCAT GGCCTTGA 1548

(2) INFORMATION FOR SEQ ID NO: 12:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1920 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: DNA (genomic)
(iii) Organism: Bacillus licheniformis
(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:421..1872
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

CGGAAGATTG GAAGTACAAA AATAAGCAAA AGATTGTCAA TCATGTCATG AGCCATGCGG 60

GAGACGGAAA AATCGTCTTA ATGCACGATA TTTATGCAAC GTTCGCAGAT GCTGCTGAAG 120

AGATTATTAA AAAGCTGAAA GCAAAAGGCT ATCAATTGGT AACTGTATCT CAGCTTGAAG 180

AAGTGAAGAA GCAGAGAGGC TATTGAATAA ATGAGTAGAA GCGCCATATC GGCGCTTTTC 240

TTTTGGAAGA AAATATAGGG AAAATGGTAC TTGTTAAAAA TTCGGAATAT TTATACAACA 300

TCATATGTTT CACATTGAAA GGGGAGGAGA ATCATGAAAC AACAAAAACG GCTTTACGCC 360

CGATTGCTGA CGCTGTTATT TGCGCTCATC TTCTTGCTGC CTCATTCTGC AGCAGCGGCG 420

GCA AAT CTT AAT GGG ACG CTG ATG CAG TAT TTT GAA TGG TAC ATG CCC 468

AAT GAC GGC CAA CAT TGG AGG CGT TTG CAA AAC GAC TCG GCA TAT TTG 516

GCT GAA CAC GGT ATT ACT GCC GTC TGG ATT CCC CCG GCA TAT AAG GGA 564

ACG AGC CAA GCG GAT GTG GGC TAC GGT GCT TAC GAC CTT TAT GAT TTA 612

GGG GAG TTT CAT CAA AAA GGG ACG GTT CGG ACA AAG TAC GGC ACA AAA 660

GGA GAG CTG CAA TCT GCG ATC AAA AGT CTT CAT TCC CGC GAC ATT AAC 708

GTT TAC GGG GAT GTG GTC ATC AAC CAC AAA GGC GGC GCT GAT GCG ACC 756

GAA GAT GTA ACC GCG GTT GAA GTC GAT CCC GCT GAC CGC AAC CGC GTA 804

ATT TCA GGA GAA CAC CTA ATT AAA GCC TGG ACA CAT TTT CAT TTT CCG 852

GGG CGC GGC AGC ACA TAC AGC GAT TTT AAA TGG CAT TGG TAC CAT TTT 900

GAC GGA ACC GAT TGG GAC GAG TCC CGA AAG CTG AAC CGC ATC TAT AAG 948

TTT CAA GGA AAG GCT TGG GAT TGG GAA GTT TCC AAT GAA AAC GGC AAC 996

TAT GAT TAT TTG ATG TAT GCC GAC ATC GAT TAT GAC CAT CCT GAT GTC 1044

GCA GCA GAA ATT AAG AGA TGG GGC ACT TGG TAT GCC AAT GAA CTG CAA 1092

TTG GAC GGT TTC CGT CTT GAT GCT GTC AAA CAC ATT AAA TTT TCT TTT 1140

TTG CGG GAT TGG GTT AAT CAT GTC AGG GAA AAA ACG GGG AAG GAA ATG 1188

TTT ACG GTA GCT GAA TAT TGG CAG AAT GAC TTG GGC GCG CTG GAA AAC 1236

TAT TTG AAC AAA ACA AAT TTT AAT CAT TCA GTG TTT GAC GTG CCG CTT 1284

CAT TAT CAG TTC CAT GCT GCA TCG ACA CAG GGA GGC GGC TAT GAT ATG 1332

AGG AAA TTG CTG AAC GGT ACG GTC GTT TCC AAG CAT CCG TTG AAA TCG 1380

GTT ACA TTT GTC GAT AAC CAT GAT ACA CAG CCG GGG CAA TCG CTT GAG 1428

TCG ACT GTC CAA ACA TGG TTT AAG CCG CTT GCT TAC GCT TTT ATT CTC 1476

ACA AGG GAA TCT GGA TAC CCT CAG GTT TTC TAC GGG GAT ATG TAC GGG 1524

ACG AAA GGA GAC TCC CAG CGC GAA ATT CCT GCC TTG AAA CAC AAA ATT 1572

GAA CCG ATC TTA AAA GCG AGA AAA CAG TAT GCG TAC GGA GCA CAG CAT 1620

GAT TAT TTC GAC CAC CAT GAC ATT GTC GGC TGG ACA AGG GAA GGC GAC 1668

AGC TCG GTT GCA AAT TCA GGT TTG GCG GCA TTA ATA ACA GAC GGA CCC 1716

GGT GGG GCA AAG CGA ATG TAT GTC GGC CGG CAA AAC GCC GGT GAG ACA 1764

TGG CAT GAC ATT ACC GGA AAC CGT TCG GAG CCG GTT GTC ATC AAT TCG 1812

GAA GGC TGG GGA GAG TTT CAC GTA AAC GGC GGG TCG GTT TCA ATT TAT 1860

GTT CAA AGA TAG AAGAGCAGAG AGGACGGATT TCCTGAAGGA AATCCGTTTT 1912

TTTATTTT 1920

(2) INFORMATION FOR SEQ ID NO: 13:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1455 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: DNA (genomic)
(iii) Organism: Bacillus sp.
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

CATCATAATG GAACAAATGG TACTATGATG CAATATTTCG AATGGTATTT GCCAAATGAC 60

GGGAATCATT GGAACAGGTT GAGGGATGAC GCAGCTAACT TAAAGAGTAA AGGGATAACA 120

```
GCTGTATGGA TCCCACCTGC ATGGAAGGGG ACTTCCCAGA ATGATGTAGG TTATGGAGCC     180
TATGATTTAT ATGATCTTGG AGAGTTTAAC CAGAAGGGGA CGGTTCGTAC AAAATATGGA     240
ACACGCAACC AGCTACAGGC TGCGGTGACC TCTTTAAAAA ATAACGGCAT TCAGGTATAT     300
GGTGATGTCG TCATGAATCA TAAAGGTGGA GCAGATGGTA CGGAAATTGT AAATGCGGTA     360
GAAGTGAATC GGAGCAACCG AAACCAGGAA ACCTCAGGAG AGTATGCAAT AGAAGCGTGG     420
ACAAAGTTTG ATTTTCCTGG AAGAGGAAAT AACCATTCCA GCTTTAAGTG GCGCTGGTAT     480
CATTTTGATG GGACAGATTG GGATCAGTCA CGCCAGCTTC AAAACAAAAT ATATAAATTC     540
AGGGGAACAG GCAAGGCCTG GGACTGGGAA GTCGATACAG AGAATGGCAA CTATGACTAT     600
CTTATGTATG CAGACGTGGA TATGGATCAC CCAGAAGTAA TACATGAACT TAGAAACTGG     660
GGAGTGTGGT ATACGAATAC ACTGAACCTT GATGGATTTA GAATAGATGC AGTGAAACAT     720
ATAAAATATA GCTTTACGAG AGATTGGCTT ACACATGTGC GTAACACCAC AGGTAAACCA     780
ATGTTTGCAG TGGCTGAGTT TTGGAAAAAT GACCTTGGTG CAATTGAAAA CTATTTGAAT     840
AAAACAAGTT GGAATCACTC GGTGTTTGAT GTTCCTCTCC ACTATAATTT GTACAATGCA     900
TCTAATAGCG GTGGTTATTA TGATATGAGA AATATTTTAA ATGGTTCTGT GGTGCAAAAA     960
CATCCAACAC ATGCCGTTAC TTTTGTTGAT AACCATGATT CTCAGCCCGG GGAAGCATTG    1020
GAATCCTTTG TTCAACAATG GTTTAAACCA CTTGCATATG CATTGGTTCT GACAAGGGAA    1080
CAAGGTTATC CTTCCGTATT TTATGGGGAT TACTACGGTA TCCCAACCCA TGGTGTTCCG    1140
GCTATGAAAT CTAAAATAGA CCCTCTTCTG CAGGCACGTC AAACTTTTGC CTATGGTACG    1200
CAGCATGATT ACTTTGATCA TCATGATATT ATCGGTTGGA CAAGAGAGGG AAATAGCTCC    1260
CATCCAAATT CAGGCCTTGC CACCATTATG TCAGATGGTC CAGGTGGTAA CAAATGGATG    1320
TATGTGGGGA AAAATAAAGC GGGACAAGTT TGGAGAGATA TTACCGGAAA TAGGACAGGC    1380
ACCGTCACAA TTAATGCAGA CGGATGGGGT AATTTCTCTG TTAATGGAGG GTCCGTTTCG    1440
GTTTGGGTGA AGCAA                                                     1455
```

(2) INFORMATION FOR SEQ ID NO: 14:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1455 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: DNA (genomic)
(iii) Organism: Bacillus sp.
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

```
CATCATAATG GGACAAATGG GACGATGATG CAATACTTTG AATGGCACTT GCCTAATGAT      60
GGGAATCACT GGAATAGATT AAGAGATGAT GCTAGTAATC TAAGAAATAG AGGTATAACC     120
GCTATTTGGA TTCCGCCTGC CTGGAAAGGG ACTTCGCAAA ATGATGTGGG GTATGGAGCC     180
TATGATCTTT ATGATTTAGG GGAATTTAAT CAAAAGGGGA CGGTTCGTAC TAAGTATGGG     240
ACACGTAGTC AATTGGAGTC TGCCATCCAT GCTTTAAAGA ATAATGGCGT TCAAGTTTAT     300
GGGGATGTAG TGATGAACCA TAAAGGAGGA GCTGATGCTA CAGAAAACGT TCTTGCTGTC     360
```

GAGGTGAATC CAAATAACCG GAATCAAGAA ATATCTGGGG ACTACACAAT TGAGGCTTGG 420

ACTAAGTTTG ATTTTCCAGG GAGGGGTAAT ACATACTCAG ACTTTAAATG GCGTTGGTAT 480

CATTTCGATG GTGTAGATTG GGATCAATCA CGACAATTCC AAAATCGTAT CTACAAATTC 540

CGAGGTGATG GTAAGGCATG GGATTGGGAA GTAGATTCGG AAAATGGAAA TTATGATTAT 600

TTAATGTATG CAGATGTAGA TATGGATCAT CCGGAGGTAG TAAATGAGCT TAGAAGATGG 660

GGAGAATGGT ATACAAATAC ATTAAATCTT GATGGATTTA GGATCGATGC GGTGAAGCAT 720

ATTAAATATA GCTTTACACG TGATTGGTTG ACCCATGTAA GAAACGCAAC GGGAAAAGAA 780

ATGTTTGCTG TTGCTGAATT TTGGAAAAAT GATTTAGGTG CCTTGGAGAA CTATTTAAAT 840

AAAACAAACT GGAATCATTC TGTCTTTGAT GTCCCCCTTC ATTATAATCT TTATAACGCG 900

TCAAATAGTG GAGGCAACTA TGACATGGCA AAACTTCTTA ATGGAACGGT TGTTCAAAAG 960

CATCCAATGC ATGCCGTAAC TTTTGTGGAT AATCACGATT CTCAACCTGG GGAATCATTA 1020

GAATCATTTG TACAAGAATG GTTTAAGCCA CTTGCTTATG CGCTTATTTT AACAAGAGAA 1080

CAAGGCTATC CCTCTGTCTT CTATGGTGAC TACTATGGAA TTCCAACACA TAGTGTCCCA 1140

GCAATGAAAG CCAAGATTGA TCCAATCTTA GAGGCGCGTC AAAATTTTGC ATATGGAACA 1200

CAACATGATT ATTTTGACCA TCATAATATA ATCGGATGGA CACGTGAAGG AAATACCACG 1260

CATCCCAATT CAGGACTTGC GACTATCATG TCGGATGGGC CAGGGGGAGA GAAATGGATG 1320

TACGTAGGGC AAAATAAAGC AGGTCAAGTT TGGCATGACA TAACTGGAAA TAAACCAGGA 1380

ACAGTTACGA TCAATGCAGA TGGATGGGCT AATTTTTCAG TAAATGGAGG ATCTGTTTCC 1440

ATTTGGGTGA AACGA 1455

(2) INFORMATION FOR SEQ ID NO: 15:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 74 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: other nucleic acid
(ix) FEATURE:
(A) NAME/ KEY: misc-feature:
(B) OTHER INFORMATION: /desc = "RSERI"
(ix) FEATURE:
(A) NAME/KEY: misc-feature
(B) LOCATION: 21-62
(D): OTHER INFORMATION: /Note= 1:97%A, 1%T, 1%C, 1%G
2:97%T, 1%A, 1%C, 1%G
3:97%C, 1%A, 1%T, 1%G
4:97%G, 1%A, 1%T, 1%C
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:
GCGTTTTGCC GGCCGACATA 3122343222 4333313344
4233423242 2122112433 43CAAACCTG AATT 74

(2) INFORMATION FOR SEQ ID NO: 16:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 122 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: other nucleic acid
(ix) FEATURE:
(A) NAME/ KEY: misc-feature:
(B) OTHER INFORMATION: /desc = "RSERII"
(ix) FEATURE:
(A) NAME/KEY: misc-feature
(B) LOCATION: 63-104
(D): OTHER INFORMATION: /Note= 1:97%A, 1%T, 1%C, 1%G
2:97%T, 1%A, 1%C, 1%G
3:97%C, 1%A, 1%T, 1%G
4:97%G, 1%A, 1%T, 1%C
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:
GCGTTTTGCC GGCCGACATA CATTCGCTTT GCCCCACCGG GTCCGTCTGT
TATTAATGCC GC31113324 1122243113 3414324234 3322333224
2331GCCGAC AATGTCATGG TG 122

(2) INFORMATION FOR SEQ ID NO: 17:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 78 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: other nucleic acid
(ix) FEATURE:
(A) NAME/ KEY: misc-feature:
(B) OTHER INFORMATION: /desc = "RSERIII"
(ix) FEATURE:
(A) NAME/KEY: misc-feature
(B) LOCATION: 19-60
(D): OTHER INFORMATION: /Note= 1:97%A, 1%T, 1%C, 1%G
2:97%T, 1%A, 1%C, 1%G
3:97%C, 1%A, 1%T, 1%G
4:97%G, 1%A, 1%T, 1%C
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:
GTCGCCTTCC CTTGTCCA43 3413112423 1244244234 1112112312
4324243233 GTACGCATAC TGTTTTCT 78

(2) INFORMATION FOR SEQ ID NO: 18:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: other nucleic acid
(ix) FEATURE:
(A) NAME/ KEY: misc-feature:
(B) OTHER INFORMATION: /desc = "FSERIII"
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:
TGGACAAGGG AAGGCGACAG 20

(2) INFORMATION FOR SEQ ID NO: 19:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 81 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: other nucleic acid
(ix) FEATURE:
(A) NAME/ KEY: misc-feature:
(B) OTHER INFORMATION: /desc = "RSERV"
(ix) FEATURE:
(A) NAME/KEY: misc-feature
(B) LOCATION: 19-60
(D): OTHER INFORMATION: /Note= 1:97%A, 1%T, 1%C, 1%G
2:97%T, 1%A, 1%C, 1%G

3:97%C, 1%A, 1%T, 1%G
4:97%G, 1%A, 1%T, 1%C
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:
TAAGATCGGT TCAATTTT42 4222311443 1441122234 3432444142
3233222342 CCCGTACATA TCCCCGTAGA A

(2) INFORMATION FOR SEQ ID NO: 20:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: other nucleic acid
(ix) FEATURE:
(A) NAME/ KEY: misc-feature:
(B) OTHER INFORMATION: /desc = "FSERV"
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:
AAAATTGAAC CGATCTTA 18

(2) INFORMATION FOR SEQ ID NO: 21:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 107 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: other nucleic acid
(ix) FEATURE:
(A) NAME/ KEY: misc-feature:
(B) OTHER INFORMATION: /desc = "FSERVII"
(ix) FEATURE:
(A) NAME/KEY: misc-feature
(B) LOCATION: 54-95
(D): OTHER INFORMATION: /Note= 1:97%A, 1%T, 1%C, 1%G
2:97%T, 1%A, 1%C, 1%G
3:97%C, 1%A, 1%T, 1%G
4:97%G, 1%A, 1%T, 1%C
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:
TTCCATGCTG CATCGACACA GGGAGGCGGC TATGATATGA GGAAATTGCT
GAA3442134 4234222331 1431233422 4111234422 13122TGTCG
ATAACCA 108

(2) INFORMATION FOR SEQ ID NO: 22:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: other nucleic acid
(ix) FEATURE:
(A) NAME/ KEY: misc-feature:
(B) OTHER INFORMATION: /desc = "RSERVII"
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:
TGTCGATGCA GCATGGAA 18

(2) INFORMATION FOR SEQ ID NO: 23:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 80 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: other nucleic acid
(ix) FEATURE:
(A) NAME/ KEY: misc-feature:

(B) OTHER INFORMATION: /desc = "FSERIX"
(ix) FEATURE:
(A) NAME/KEY: misc-feature
(B) LOCATION: 21-62
(D): OTHER INFORMATION: /Note= 1:97%A, 1%T, 1%C, 1%G
2:97%T, 1%A, 1%C, 1%G
3:97%C, 1%A, 1%T, 1%G
4:97%G, 1%A, 1%T, 1%C
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:
GTCCAAACAT GGTTTAAGCC 4322432213 4322221223 2313114441
1232441213 33TCAGGTTT TCTACGGGGA 80

(2) INFORMATION FOR SEQ ID NO: 24:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: other nucleic acid
(ix) FEATURE:
(A) NAME/ KEY: misc-feature:
(B) OTHER INFORMATION: /desc = "RSERIX"
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:
GGCTTAAACC ATGTTTGGAC 20

(2) INFORMATION FOR SEQ ID NO: 26:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: other nucleic acid
(ix) FEATURE:
(A) NAME/ KEY: misc-feature:
(B) OTHER INFORMATION: /desc = "Primer 1B"
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:
CGATTGCTGA CGCTGTTATT TGCG 24

(2) INFORMATION FOR SEQ ID NO: 27:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: other nucleic acid
(ix) FEATURE:
(A) NAME/ KEY: misc-feature:
(B) OTHER INFORMATION: /desc = "Primer #63"
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:
CTATCTTTGA ACATAAATTG AAACC 25

(2) INFORMATION FOR SEQ ID NO: 28:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: other nucleic acid
(ix) FEATURE:
(A) NAME/ KEY: misc-feature:
(B) OTHER INFORMATION: /desc = "forward Primer1"
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:
gacctgcagt caggcaacta 20

(2) INFORMATION FOR SEQ ID NO: 29:
(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: other nucleic acid
(ix) FEATURE:
(A) NAME/ KEY: misc-feature:
(B) OTHER INFORMATION: /desc = "reverse primer 1"
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

tagagtcgac ctgcaggcat 20

(2) INFORMATION FOR SEQ ID NO: 30:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: other nucleic acid
(ix) FEATURE:
(A) NAME/ KEY: misc-feature:
(B) OTHER INFORMATION: /desc = "forward primer 2"
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

gacctgcagt caggcaacta 20

(2) INFORMATION FOR SEQ ID NO: 31:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: other nucleic acid
(ix) FEATURE:
(A) NAME/ KEY: misc-feature:
(B) OTHER INFORMATION: /desc = "reverse primer 2"
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

tagagtcgac ctgcaggcat 20

(2) INFORMATION FOR SEQ ID NO: 32:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2084 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: DNA (genomic)
(iii)Organism: Bacillus amyloliquefaciens
(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:343..1794
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

GCCCCGCACA TACGAAAAGA CTGGCTGAAA ACATTGAGCC TTTGATGACT GATGATTTGG 60

CTGAAGAAGT GGATCGATTG TTTGAGAAAA GAAGAAGACC ATAAAAATAC CTTGTCTGTC 120

ATCAGACAGG GTATTTTTTA TGCTGTCCAG ACTGTCCGCT GTGTAAAAAT AAGGAATAAA 180

GGGGGGTTGT TATTATTTTA CTGATATGTA AAATATAATT TGTATAAGAA AATGAGAGGG 240

AGAGGAAACA TGATTCAAAA ACGAAAGCGG ACAGTTTCGT TCAGACTTGT GCTTATGTGC 300

ACGCTGTTAT TTGTCAGTTT GCCGATTACA AAAACATCAG CC GTA AAT GGC ACG 354

CTG ATG CAG TAT TTT GAA TGG TAT ACG CCG AAC GAC GGC CAG CAT TGG 402

AAA CGA TTG CAG AAT GAT GCG GAA CAT TTA TCG GAT ATC GGA ATC ACT 450

GCC GTC TGG ATT CCT CCC GCA TAC AAA GGA TTG AGC CAA TCC GAT AAC 498

GGA TAC GGA CCT TAT GAT TTG TAT GAT TTA GGA GAA TTC CAG CAA AAA 546

GGG ACG GTC AGA ACG AAA TAC GGC ACA AAA TCA GAG CTT CAA GAT GCG 594

ATC GGC TCA CTG CAT TCC CGG AAC GTC CAA GTA TAC GGA GAT GTG GTT 642

TTG AAT CAT AAG GCT GGT GCT GAT GCA ACA GAA GAT GTA ACT GCC GTC 690

GAA GTC AAT CCG GCC AAT AGA AAT CAG GAA ACT TCG GAG GAA TAT CAA 738

ATC AAA GCG TGG ACG GAT TTT CGT TTT CCG GGC CGT GGA AAC ACG TAC 786

AGT GAT TTT AAA TGG CAT TGG TAT CAT TTC GAC GGA GCG GAC TGG GAT 834

GAA TCC CGG AAG ATC AGC CGC ATC TTT AAG TTT CGT GGG GAA GGA AAA 882

GCG TGG GAT TGG GAA GTA TCA AGT GAA AAC GGC AAC TAT GAC TAT TTA 930

ATG TAT GCT GAT GTT GAC TAC GAC CAC CCT GAT GTC GTG GCA GAG ACA 978

AAA AAA TGG GGT ATC TGG TAT GCG AAT GAA CTG TCA TTA GAC GGC TTC 1026

CGT ATT GAT GCC GCC AAA CAT ATT AAA TTT TCA TTT CTG CGT GAT TGG 1074

GTT CAG GCG GTC AGA CAG GCG ACG GGA AAA GAA ATG TTT ACG GTT GCG 1122

GAG TAT TGG CAG AAT AAT GCC GGG AAA CTC GAA AAC TAC TTG AAT AAA 1170

ACA AGC TTT AAT CAA TCC GTG TTT GAT GTT CCG CTT CAT TTC AAT TTA 1218

CAG GCG GCT TCC TCA CAA GGA GGC GGA TAT GAT ATG AGG CGT TTG CTG 1266

GAC GGT ACC GTT GTG TCC AGG CAT CCG GAA AAG GCG GTT ACA TTT GTT 1314

GAA AAT CAT GAC ACA CAG CCG GGA CAG TCA TTG GAA TCG ACA GTC CAA 1362

ACT TGG TTT AAA CCG CTT GCA TAC GCC TTT ATT TTG ACA AGA GAA TCC 1410

GGT TAT CCT CAG GTG TTC TAT GGG GAT ATG TAC GGG ACA AAA GGG ACA 1458

TCG CCA AAG GAA ATT CCC TCA CTG AAA GAT AAT ATA GAG CCG ATT TTA 1506

AAA GCG CGT AAG GAG TAC GCA TAC GGG CCC CAG CAC GAT TAT ATT GAC 1554

CAC CCG GAT GTG ATC GGA TGG ACG AGG GAA GGT GAC AGC TCC GCC GCC 1602

AAA TCA GGT TTG GCC GCT TTA ATC ACG GAC GGA CCC GGC GGA TCA AAG 1650

CGG ATG TAT GCC GGC CTG AAA AAT GCC GGC GAG ACA TGG TAT GAC ATA 1698

ACG GGC AAC CGT TCA GAT ACT GTA AAA ATC GGA TCT GAC GGC TGG GGA 1746

GAG TTT CAT GTA AAC GAT GGG TCC GTC TCC ATT TAT GTT CAG AAA TAA 1794

GGTAATAAAA AAACACCTCC AAGCTGAGTG CGGGTATCAG CTTGGAGGTG CGTTTATTTT 1854

TTCAGCCGTA TGACAAGGTC GGCATCAGGT GTGACAAATA CGGTATGCTG GCTGTCATAG 1914

GTGACAAATC CGGGTTTTGC GCCGTTTGGC TTTTTCACAT GTCTGATTTT TGTATAATCA 1974

ACAGGCACGG AGCCGGAATC TTTCGCCTTG GAAAAATAAG CGGCGATCGT AGCTGCTTCC 2034

AATATGGATT GTTCATCGGG ATCGCTGCTT TTAATCACAA CGTGGGATCC 2084

INTERNATIONAL SEARCH REPORT

International application No.
PCT/DK 99/00628

A. CLASSIFICATION OF SUBJECT MATTER

IPC7: C12N 9/28

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC7: C12N, C11D

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 9510603 A1 (NOVO NORDISK A/S), 20 April 1995 (20.04.95), see claim 14	1-16,18-41
A		17
	--	
X	WO 9623873 A1 (NOVO NORDISK A/S), 8 August 1996 (08.08.96), page 23, line 24 and forward	41
A		1-39
	--	
A	WO 9743424 A1 (GENENCOR INTERNATIONAL, INC.), 20 November 1997 (20.11.97), page 12, lines 4-24	1-41
	--	

☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" erlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
21 March 2000	06 -04- 2000

Name and mailing address of the ISA/
Swedish Patent Office
Box 5055, S-102 42 STOCKHOLM
Facsimile No. +46 8 666 02 86

Authorized officer
Carolina Palmcrantz/EÖ
Telephone No. +46 8 782 25 00

INTERNATIONAL SEARCH REPORT

International application No.
PCT/DK 99/00628

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: 1
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

see next sheet

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/DK 99/00628

The wording "solvent exposed amino acid residues" of claim 1 may include other (unknown) residues in addition to those given in the present application (c.f. the specification page 16, lines 10-13). Therefore, the search has been incomplete and restricted to those residues specified in the application (c.f. PCT, Article 6).

INTERNATIONAL SEARCH REPORT

Information on patent family members

02/12/99

International application No.

PCT/DK 99/00628

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9510603 A1	20/04/95	AU 7807494 A	04/05/95
		BR 9407767 A	18/03/97
		CA 2173329 A	20/04/95
		CN 1134725 A	30/10/96
		EP 0722490 A	24/07/96
		FI 961524 A	30/05/96
		JP 9503916 T	22/04/97
		US 5753460 A	19/05/98
		US 5801043 A	01/09/98
WO 9623873 A1	08/08/96	AU 4483396 A	21/08/96
		BR 9607735 A	14/07/98
		CA 2211405 A	08/08/96
		CN 1172500 A	04/02/98
		EP 0815208 A	07/01/98
		JP 11503003 T	23/03/99
WO 9743424 A1	20/11/97	AU 2996997 A	05/12/97
		EP 0927259 A	07/07/99
		US 5763385 A	09/06/98